



PCT

ORGANISATION MONDIALE DE LA PROPRIETE INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

| | | | |
|--|--|--|---|
| (51) Classification internationale des brevets ⁴ : C07K 7/10, 7/06, C12N 15/00 G01N 33/569, A61K 39/21, 37/02 | | A1 | (11) Numéro de publication internationale: WO 88/05440 (43) Date de publication internationale: 28 juillet 1988 (28.07.88) |
| (21) Numéro de la demande internationale: | PCT/FR88/00025 | (72) Inventeurs; et | |
| (22) Date de dépôt international: | 15 janvier 1988 (15.01.88) | (75) Inventeurs/Déposants (<i>US seulement</i>): ALIZON, Marc [FR/ FR]; 71, rue du Cardinal-Lemoine, F-75005 Paris (FR). MONTAGNIER, Luc [FR/FR]; 21, rue de Malabry, F- 92350 Le-Plessis-Robinson (FR). GUETARD, Denise [FR/ FR]; 4 B, rue Anselme-Payen, F-75015 Paris (FR). CLA- VEL, François [FR/US]; 12103 Postree Drive, Rockville, MD 20852 (US). SONIGO, Pierre [FR/FR]; 23, rue Guten- berg, F-75015 Paris (FR). GUYADER, Mireille [FR/FR]; 68, rue Laugier, F-75017 Paris (FR). TIOLLAIS, Pierre [FR/ FR]; 16, rue de la Glacière, F-75013 Paris (FR). CHAKRA- BARTI, Lisa [FR/FR]; 16, rue des 3 Portes, F-75005 Paris (FR). DESROSIERS, Ronald [US/US]; 13 Causeway Street, Udson, MA 01749 (US). | |
| (31) Numéros des demandes prioritaires: | 003,764 87/01739 87/05398 | (74) Mandataires: GUTMANN Ernest etc.; S.C. Ernest Gutmann - Yves Plasserand, 67, boulevard Haussmann, F-75008 Paris (FR). | |
| (32) Dates de priorité: | 16 janvier 1987 (16.01.87) 11 février 1987 (11.02.87) 15 avril 1987 (15.04.87) | (81) Etats désignés: AU, DK, JP, KR, US. | |
| (33) Pays de priorité: | US FR FR | Publiée <i>Avec rapport de recherche internationale. Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.</i> | |
| (60) Brevet ou demande principal(e) | | | |
| (63) Apparenté(e) par continuation | | | |
| US | 013,477 (CIP) | | |
| Déposée le | 11 février 1987 (11.02.87) | | |
| (71) Déposant (<i>pour tous les Etats désignés sauf US</i>): INSTITUT PASTEUR [FR/FR]; 25-28, rue du Dr.-Roux, F-75015 Paris (FR). | | | |

(54) Title: PEPTIDES HAVING IMMUNOLOGICAL PROPERTIES 2-HJV-2

(54) Titre: PEPTIDES AYANT DES PROPRIETES IMMUNOLOGIQUES 2-HIV-2

(57) Abstract

Peptides having immunological properties in common with those of the peptidic skeleton of peptides of viruses of the family HIV-2, particularly the envelope glycoprotein of HIV-2, characterized in that they have also a peptidic structure in common with the peptidic skeleton of peptides of SIV, particularly the envelope glycoprotein of SIV. The invention also relates to diagnosis compositions capable of detecting an infection due to HIV-2 and to vaccine compositions.

(57) Abrégé

Peptides ayant des propriétés immunologiques en commun avec celles de l'ossature peptidique des peptides des virus de la classe HIV-2, notamment de la glycoprotéine d'enveloppe de HIV-2, caractérisés en ce qu'ils ont également une structure peptidique en commun avec l'ossature peptidique des peptides de SIV, notamment de la glycoprotéine d'enveloppe de SIV. L'invention concerne des compositions de diagnostic capable de détecter une infection due à HIV-2 et des compositions de vaccin.

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

| | | | | | |
|----|-----------------------------------|----|---|----|-----------------------|
| AT | Autriche | FR | France | ML | Mali |
| AU | Australie | GA | Gabon | MR | Mauritanie |
| BB | Barbade | GB | Royaume-Uni | MW | Malawi |
| BE | Belgique | HU | Hongrie | NL | Pays-Bas |
| BG | Bulgarie | IT | Italie | NO | Norvège |
| BJ | Bénin | JP | Japon | RO | Roumanie |
| BR | Brésil | KP | République populaire démocratique de Corée | SD | Soudan |
| CF | République Centrafricaine | KR | République de Corée | SE | Suède |
| CG | Congo | LI | Liechtenstein | SN | Sénégal |
| CH | Suisse | LK | Sri Lanka | SU | Union soviétique |
| CM | Cameroun | LU | Luxembourg | TD | Tchad |
| DE | Allemagne, République fédérale d' | MC | Monaco | TG | Togo |
| DK | Danemark | MG | Madagascar | US | Etats-Unis d'Amérique |

5

Peptides ayant des propriétés immunologiques de HIV-2

La présente invention est relative à des peptides ayant des propriétés immunologiques, le cas échéant immunogènes, en commun avec des antigènes susceptibles d'être obtenus sous une forme purifiée, à partir de virus capables de provoquer des lymphadénopathies susceptibles d'être relayées ensuite par le syndrome d'immunodéficience acquise (SIDA) chez l'homme.

L'invention concerne en particulier des peptides antigéniques susceptibles d'être reconnus par des anticorps induits chez l'homme par des virus désignés par l'abréviation HIV, selon la nomenclature définie dans NATURE. Elle concerne également des peptides ayant des propriétés immunogènes ou susceptibles d'être rendus immunogènes *in vivo*, cette immunogénicité étant susceptible de se manifester par l'induction *in vivo* d'anticorps reconnaissant des antigènes caractéristiques des virus HIV-2 et même, au moins en ce qui concerne certains de ces peptides, des antigènes issus de HIV-1.

L'invention concerne en outre des applications de ces peptides à la fabrication de compositions pour le diagnostic *in vitro* chez l'homme de potentialité de certaines formes du SIDA et, en ce qui concerne certains d'entre eux, à la production de compositions immunogènes et de compositions vaccinantes contre les rétrovirus HIV.

De même l'invention concerne les applications aux mêmes fins des anticorps susceptibles d'être induits

in vivo par les peptides immunogènes ou rendus immuno-gènes et, pour certains de ces anticorps, leurs applications à la production de principes actifs de médicaments contre ces SIDAS humains.

5 L'invention concerne également la mise en oeuvre de certains de ces peptides dans des procédés pour le diagnostic in vitro chez l'homme de certaines formes du SIDA, ainsi que leur application à la constitution de trousse ou "kits" de diagnostic.

10 Un premier rétrovirus dénommé LAV-1 ou HIV-1 a été isolé et décrit dans la demande de brevet GB.83/24.800 et une demande EP.84/401.834 du 14/09/84. Ce virus a également été décrit par F.Barre Sinoussi et al. dans Science, 220 n° 45-99, 20 pages 868-871.

15 Des variants de ce virus HIV-1 désignés par LAV ELI et LAV MAL, ont également été isolés, caractérisés et décrits dans la demande de brevet EP.84/-401.834.

20 Les virus HIV-1 et leurs variants possèdent les propriétés suivantes :

- ils ont pour cibles préférentielles les cellules Leu3 (ou lymphocytes T4) humaines et leurs cellules dérivées "immortalisées".

25 - ils ont une activité transcriptase inverse nécessitant la présence d'ions Mg²⁺ et présentent une forte activité pour le poly(adenylate-oligo-deoxythymidylase) poly(A)-oligo(dT)12-18

- ils ont une densité de 1,16 à 1,17 sur gradient de sucre,

30 - ils ont un diamètre moyen de 139 nanomètres et un noyau de diamètre moyen de 41 nanomètres,

- les lysats de ces virus contiennent une protéine p25 (protéine du noyau) qui ne croise pas immunologiquement avec la protéine p24 de HTLV-1,

35 - ils contiennent une protéine p42 appartenant à leur enveloppe,

- ils contiennent également une glycoprotéine d'enveloppe gp110 d'un poids moléculaire de 110.000.

L'isolement et la caractérisation de rétro-virus appartenant à une classe distincte et n'ayant qu'une parenté immunologique réduite avec les précédents, ont été décrits dans la demande de brevet européen n° 87/400.151.4. Ces rétrovirus qui ont été regroupés sous la désignation HIV-2, ont été isolés chez plusieurs malades africains présentant des symptômes d'une lymphadénopathie ou d'un SIDA.

Les rétrovirus du type HIV-2 comme les rétro-virus du type HIV-1, se caractérisent par un tropisme pour les lymphocytes T4 humains et par un effet cytopathogène à l'égard de ces lymphocytes, lorsqu'ils s'y multiplient, pour alors causer soit des poly-adénopathies généralisées et persistantes, soit un SIDA.

Plus généralement les rétrovirus purifiés par HIV-2 possèdent en général les propriétés suivantes :

- la cible préférentielle des rétrovirus HIV-2 est constituée par les cellules Leu3 (ou lymphocytes T4) humaines et pour des cellules "immortalisées" dérivées de ces lymphocytes T4 ;
- ils sont cytotoxiques pour les lymphocytes T4 humains
- ils ont une activité de transcriptase inverse nécessitant la présence d'ions Mg^{2+} et présentant une forte activité pour le poly(adénylate-oligodéoxythymidylase) (poly(A)-oligo(dT) 12-18) ;
- ils ont une densité de 1,16 dans un gradient de sucre ;
- ils ont un diamètre moyen de 140 nanomètres et un noyau ayant un diamètre moyen de 41 nanomètres ;
- ils peuvent être cultivés dans des lignées permanentes du type HUT ou exprimant la protéine T4 ;
- ils ne sont pas infectieux pour les lymphocytes T8 ;
- les lysats de ces virus contiennent une protéine p26

- qui ne croise pas immunologiquement avec la protéine p24 du virus HTLV-I ou du virus HTLV-II ;
- ces lysats contiennent en outre une protéine p16 qui n'est pas reconnue immunologiquement par la protéine p19 de HTLV-I ou de HTLV-II dans des essais de radioimmuno-précipitation ;
- 5 - ils contiennent en outre une glycoprotéine d'enveloppe ayant un poids moléculaire de l'ordre de 130.000-140.000 qui ne croise pas immunologiquement avec la gp110 des HIV-1, mais qui en revanche croise immunologiquement avec la glycoprotéine d'enveloppe gp140 de STLV-III (virus isolé chez le singe) ;
- 10 - ces lysats contiennent encore des antigènes marquables par la ³⁵S-cystéine, dont les poids moléculaires s'étagent entre 32.000 et 42.000-45.000 : ils comprennent notamment un antigène ayant un poids moléculaire de l'ordre de 36.000 et un antigène ayant un poids moléculaire de l'ordre de 42.000, l'un de ces antigènes (p36 et p42) constituant vraisemblablement une glycoprotéine trans-membranaire du virus HIV-2 ;
- 15 - l'ARN génomique des HIV-2 n'hybride pas avec l'ARN génomique de HIV-1 dans des conditions stringentes ;
- dans des conditions non stringentes, l'ARN génomique de HIV-2 n'hybride, ni avec le gène env et le LTR qui le jouxte, de HIV-1, ni avec des séquences de la région pol du génome de HIV-1 ;
- 20 - dans des conditions non stringentes, il hybride faiblement avec des séquences de nucléotides de la région de HIV-1.
- 25 30 Un autre rétrovirus dénommé SIV-1, cette dénomination remplaçant la dénomination antérieurement connue STLV III, a été isolé chez le singe macaque rhésus. (M.D.Daniel et al. Science 228, 1201 (1985) N.L.Letwin et al, Science 230, 71 (1985) sous l'appellation "STLV-III Mac").
- 35

Un autre rétrovirus, désigné "STLV-III_{AGM}", (ou SIV_{AGM}) a été isolé chez des singes verts sauvages. Mais, contrairement au virus présent chez le singe macaque rhésus, la présence de "STLV-III_{AGM}" ne semble pas induire une maladie du type SIDA chez le singe vert d'Afrique.

Une souche du rétrovirus SIV-1mac a été déposée à la CNCM le 7 Février 1986 sous le n° I-521. Des études ont montré que le rétrovirus SIV-1 comporte certaines protéines possédant une certaine parenté immunologique avec des protéines ou glycoprotéines structurales susceptibles d'être obtenues dans des conditions analogues, à partir de HIV-2. Ce rétrovirus SIV-1, dont on a constaté le caractère infectieux chez les singes, avait été désigné par STL VIII par les chercheurs qui l'ont isolé (références bibliographiques précitées).

Pour la commodité du langage, ces virus ne seront plus désignés dans ce qui suit que par l'expression SIV (l'expression SIV est l'abréviation anglaise de "Simian Immunodeficiency Virus" (virus d'immunodéficience du singe)) éventuellement suivi d'une abréviation désignant l'espèce de singe dont ils sont issus, par exemple, MAC (ou mac) pour le macaque ou AGM pour le singe vert d'Afrique (abréviation de "African Green Monkey").

En mettant en oeuvre les mêmes techniques que celles rappelées plus haut, il a été constaté que l'on pouvait également obtenir à partir de SIV-1mac :

- une protéine principale du noyau p27, ayant un poids moléculaire de l'ordre de 27 kilodaltons,
- une glycoprotéine majeure d'enveloppe, gp140,
- une protéine vraisemblablement transmembranaire p32, qui n'est guère observée en RIPA lorsque le virus a au préalable été marqué par la ³⁵S-cystéine, mais qui peut

être observée dans les essais d'immunoempreintes (Western blots), sous forme de bandes larges.

Des études plus précises ont été réalisées en ce qui concerne les précédents virus HIV-2 et SIV. La poursuite de l'étude des rétrovirus HIV-2 a également conduit à l'obtention de séquences d'ADN complémentaires (ADNC) des ARNs de leurs génomes. La séquence nucléotidique complète de l'ADNC d'un rétrovirus représentatif de la classe HIV-2 (HIV-2 ROD) a été déposée le 21/02/-1986 à la CNCM sous le n° I-522, sous le nom de référence LAV-II ROD).

Cette séquence nucléotidique et les phases de lecture ouverte qu'elle contient sont indiqués à la figure 1 A.

En outre, la poursuite de l'étude d'autres rétrovirus a également permis d'aboutir à l'obtention de leurs séquences nucléotidiques complètes. Il en est en particulier ainsi de l'ADNC dérivé de l'ARN génomique de SIV.

Le clonage et le séquençage du virus SIV-1mac qui ont permis l'obtention de sa séquence nucléotidique ont été réalisés dans les conditions suivantes :

L'ADN de cellules HUT 78 infectées par le virus SIV (isolat STLV-III mac 142-83 décrit par Daniel et al. (1985) *Science*, 228, p. 1201-1204, digéré partiellement par l'enzyme de restriction Sau3A a été cloné au site BamHI du bactériophage vecteur Lambda ELBL3 pour constituer une banque génomique. Les 2 millions de phages recombinants de la banque génomique ainsi constituée ont été criblés in situ en conditions de sécurité P3, à l'aide de séquences du virus HIV2 provenant des clones lambda-ROD4, lambda-ROD35 et E2 (Clavel et al. (1986-*Nature*, 324, p. 691.) et nick-translatées.

L'hybridation a été réalisée en 5xSSC à 50°C et les lavages en 2xSSC à 50°C. Un seul clone contenant

l'ensemble des séquences virales a été obtenu. Ce clone est désigné par lambda-SIV-1. L'insérat du phage lambda-SIV-1 mesure 16,5 kb au total et comprend un provirus intégré auquel manquent seulement les 250 premières bases du LTR gauche, alors que le LTR droit est complet.

Le provirus intégré a été séquencé par la méthode des didéoxynucléotides après sous-clonage de fragments aléatoires dans le phage M13mp8. 300 sous-clones ont été analysés.

Des fragments d'ADNc provenant du clone Lambda SIV-1 insérés dans des plasmides pSIV-1.1 et pSIV-1.2 ont été déposés à la CNCM le 15 Avril 1987, sous les numéros I-658 (pSIV-1.1) et I-659 (pSIV-1.2).

Les résultats ont été mentionnés dans les figures décrites ci-après.

La figure 1B représente la séquence nucléotidique du génome viral de SIV et les séquences qui en sont déduites pour les protéines virales correspondant aux produits des gènes gag, pol, env, Q, X, R, tat, art, F.

Les figures 3 à 11 et la figure 1C représentent les comparaisons des produits théoriques des gènes viraux et des LTR entre HIV2 et SIVmac. (λ SIV-1).

L'invention concerne de plus les fragments d'ADNc déduits de l'ADNc issu du génome entier de SIV-1, ces fragments contenant une ou plusieurs séquences issues de la séquence complète d'ADNc et qui codent pour des peptides intéressants de l'invention. Ces séquences sont indiquées à la figure 1B et, à la figure 1C pour ce qui a trait à la séquence LTR du virus,

Les séquences nucléiques de l'ADNc de SIV ont été placées en correspondance avec les séquences nucléiques du virus HIV-2 ROD pour ce qui concerne la séquence LTR (figure 1C). Cette présentation que l'on retrouve pour le génome entier en rapprochant la figure 1B

des figures 3 à 11 permet de repérer ou de déduire les acides nucléiques ayant des éléments de structure essentiels communs aux deux virus.

L'invention concerne naturellement aussi l'utilisation des cADNs issus de SIV ou de leurs fragments (ou de recombinants les contenant) en tant que sondes, pour le diagnostic de la présence ou non de virus HIV-2 dans des échantillons de sérum ou d'autres liquides ou tissus biologiques obtenus à partir de patients suspectés d'être porteurs du virus HIV-2. Ces sondes sont de préférence marquées également (marqueurs radio-actifs, enzymatiques, fluorescents, etc.). Des sondes particulièrement intéressantes pour la mise en oeuvre du procédé de diagnostic du virus HIV-2 ou d'un variant de HIV-2 peuvent être caractérisées en ce qu'elles comprennent la totalité ou une fraction de l'ADNc complémentaire du génome du virus SIV ou encore notamment les fragments recombinants contenus dans divers clones.

Les sondes mises en oeuvre dans ce procédé de diagnostic du virus HIV-2 et dans les kits de diagnostic ne sont en aucune façon réduites aux sondes décrites précédemment. Elles comprennent au contraire toutes les séquences nucléotidiques issues du génome du virus SIV, d'un variant de SIV ou d'un virus proche par sa structure, dès lors qu'elles permettent la détection dans des fluides biologiques de personnes susceptibles de développer un SIDA, d'anticorps dirigés contre un HIV-2 ou d'un virus qui en est proche.

La détection peut être réalisée de toutes façons en soi connues. Elle peut comprendre une mise en contact de ces sondes soit avec les acides nucléiques obtenus à partir des cellules contenues dans ces sérum ou autres milieux biologiques, par exemple liquides céphalo-rachidiens, salives, etc... Elle peut aussi

comprendre une mise en contact de ces sondes avec ces milieux eux-mêmes dès lors que leurs acides nucléiques ont été rendus accessibles à l'hybridation avec ces sondes, et ce dans des conditions permettant l'hybridation entre ces sondes et ces acides nucléiques. L'é-
5 tape finale du diagnostic in vitro comprend alors la détection de l'hybridation éventuellement produite. Le susdit diagnostic mettant en jeu des réactions d'hy-
bridation peut également être réalisé à l'aide de mé-
10 langes de sondes respectivement originaires d'un HIV-2 et d'un SIV-1 ou d'un HIV-1, d'un HIV-2 et d'un SIV, dès lors qu'il n'est pas nécessaire de faire une différence entre le type de virus recherché.

D'une façon générale, le procédé de diagnostic de la présence ou non du virus HIV-2 ou d'un variant dans des échantillons de sérum ou d'autres liquides ou tissus, obtenus à partir de patients suspectés d'être porteurs du virus HIV-2 comprend les étapes suivantes :

1/ au moins une étape d'hybridation conduite dans des conditions stringentes, par mise en contact de l'ADN de cellules de l'échantillon du patient suspect avec l'une des susdites sondes marquées sur une membrane appropriée,

2/ le lavage de ladite membrane avec une so-
lution assurant la conservation de ces conditions strin-
gentes de l'hybridation,

3/ la détection de la présence ou non du virus HIV-2 par une méthode d'immunodétection.

Dans un autre mode de réalisation préféré du procédé selon l'invention l'hybridation précitée est conduite dans des conditions non stringentes et le la-
vage de la membrane est réalisé dans des conditions adaptées à celles de l'hybridation.

Il va de soi que l'invention concerne les
35 acides nucléiques correspondant à des séquences placées

en des régions analogues de variants de SIV ainsi que tous les acides nucléiques dont les modifications résulteraient de la mise à profit de la dégénérescence du code génétique.

5 Les études comparatives qui ont aussi permis d'aboutir à des résultats relatifs aux protéines de noyau (core), ci-après dénommées "protéines gag" et aux protéines d'enveloppes, ci-après dénommées "protéines env", ont également été rapportés dans la demande de brevet européen n° 87/400.151.4, déjà citée. Ces résultats montrent que les protéines du noyau (protéines gag) dans HIV-2 présentent des différences moins accentuées par rapport à celles des virus HIV-1, que les protéines d'enveloppe (protéines env). Globalement les protéines env dans HIV-2 se sont révélées présenter des parentés immunologiques extrêmement faibles, sinon inexistantes, avec les protéines env correspondantes des virus HIV-1.

20 Au contraire des études comparatives effectuées entre les structures des séquences d'ADNc des virus HIV-2 et SIV permettent de mettre en évidence certaines caractéristiques communes qui apparaissent au niveau des protéines.

25 Globalement, les protéines de HIV-2 et de SIV-1 montrent des parentés immunologiques importantes.

30 La glycoprotéine majeure d'enveloppe de HIV-2 s'est révélée être plus proche immunologiquement de la glycoprotéine majeure d'enveloppe de SIV que de la glycoprotéine majeure d'enveloppe de HIV-1.

35 Ces constatations s'imposent non seulement au niveau des poids moléculaires : 130-140 kilodaltons pour les glycoprotéines majeures de HIV-2 et de SIV contre environ 110 pour la glycoprotéine majeure d'enveloppe de HIV-1, mais aussi au niveau des propriétés immunologiques, puisque des sérums prélevés à partir de malades

infectés par HIV-2, et plus particulièrement des anticorps formés contre la gp140 de HIV-2 reconnaissent la gp140 de SIV-1mac, alors que dans des essais semblables les mêmes sérum et les mêmes anticorps de HIV-2 ne reconnaissent pas la gp110 de HIV-1. Mais les sérum anti-HIV-1 qui n'ont jamais réagi avec la gp140 de HIV-2 précipitent une protéine de 26 Kdal marquée par la ³⁵S-cystéine, contenue dans les extraits de HIV-2.

La protéine majeure du noyau (core) de HIV-2 semble présenter un poids moléculaire moyen (environ 10 26.000) intermédiaire entre celui de la p25 de HIV-1 et la p27 de SIV.

Ces observations résultent des essais réalisés avec des extraits viraux obtenus à partir du HIV-2 isolé 15 à partir de l'un des patients susmentionnés. Des résultats similaires ont été obtenus avec des extraits viraux du HIV-2 isolé à partir du second patient.

Des études plus poussées ont conduit les inventeurs à reconnaître une première classe de peptides 20 ayant des séquences d'aminoacides soit identiques, soit proches de séquences contenues à l'intérieur des structures des protéines gag et env de HIV-2 ou de SIV voire de HIV-1. Ces peptides sont notamment applicables au diagnostic d'une infection chez l'homme par le virus 25 HIV-2 ou de l'un de ses variants.

A cet égard la présente invention concerne également des procédés et des compositions de diagnostic pour la détection in vitro d'anticorps dirigés contre un virus HIV-2 ou de ses variants, plus particulièrement dans des échantillons biologiques, notamment des sérum de patients ayant subi une infection par le virus HIV-2, certains de ces peptides permettant une discrimination 30 particulièrement poussée entre les infections dues à des virus HIV-2 et à des virus HIV-1.

35 Ces études poussées ont également conduit à la

possibilité de synthétiser des peptides immunogènes ou susceptibles d'être rendus immunogènes, présentant des caractéristiques de structures leur permettant d'induire in vivo la production d'anticorps susceptibles de reconnaître des protéines env à la fois dans HIV-1 et dans HIV-2 et, au moins pour certains de ces peptides, de se fixer tant sur des virus HIV-1 que sur des virus HIV-2, plus particulièrement aux fins de les neutraliser. L'utilisation de ces derniers types de peptides est donc particulièrement indiquée pour la production de principes actifs de vaccins contre les virus HIV, donc contre le SIDA.

Pour désigner ci-après les résidus d'amino-acides entrant dans la constitution des peptides selon l'invention, on aura recours, pour ceux des acides aminés ayant une signification univoque à la nomenclature internationale désignant chaque acide aminé naturel par une lettre unique (lettre majuscule) selon le tableau des correspondances qui suit :

| | | |
|----|---|--------------------|
| 20 | M | Méthionine |
| | L | Leucine |
| | I | Isoleucine |
| | V | Valine |
| | F | Phenylalanine |
| 25 | S | Sérine |
| | P | Proline |
| | T | Thréonine |
| | A | Alanine |
| | Y | Tyrosine |
| 30 | H | Histidine |
| | Q | Glutamine |
| | N | Asparagine |
| | K | Lysine |
| | D | Acide Aspartique |
| 35 | E | Acide glutaminique |

| | |
|---|-------------|
| C | Cystéine |
| W | Tryptophane |
| R | Arginine |
| G | Glycine |

5 Lorsqu'un acide aminé pourra, en raison de sa position au sein de la chaîne d'aminoacides caractéristique d'un peptide déterminé, prendre plusieurs significations, il pourra soit être désigné par un tiret "--", si sa signification peut être quelconque, soit par 10 une lettre minuscule lorsque cet aminoacide pourra présenter un nombre limité de significations préférées, ce nombre étant cependant toujours supérieur à 1. Dans ce dernier cas, les significations possibles de cette lettre minuscule seront toujours précisées en rapport 15 avec le peptide auquel il appartient.

Afin de faciliter la lecture, ces peptides seront désignés par une abréviation env ou gag suivie d'un indice numérique, par référence à des séquences d'aminoacides contenues, selon le cas, soit dans les 20 protéines env soit dans les protéines gag de certains HIV-1, HIV-2 ou SIV. Il y sera encore fait référence dans ce qui suit.

Enfin dans les définitions qui suivent

- les groupes X représentent soit un groupe NH₂ libre ou amidé, notamment par un ou deux groupes alcoyle comprenant de 1 à 5 atomes de carbone, soit un groupe peptidique comprenant de 1 à 5 aminoacides, dont l'aminoacide N-terminal présente lui-même un groupe NH₂ libre ou amidé comme précédemment indiqué, et
- les groupes Z représentent, soit un groupe -OH libre ou alcoxyle et contenant alors un groupe alcoyle comprenant de 1 à 5 atomes de carbone, soit un groupe peptidique comprenant de 1 à 5 aminoacides, dont l'aminoacide C-terminal présente lui-même un groupe -OH libre ou alcoxyle, comme précédemment indiqué, les

groupes de 1 à 5 acides aminés le cas échéant contenus dans X ou Z ou dans les deux à la fois étant tels, que leur présence n'est pas incompatible avec la préservation pour l'essentiel des propriétés immunologiques, le cas échéant immunogènes, des peptides qui en sont dépourvus.

Les peptides selon l'invention, qui ont en commun des propriétés immunologiques avec des antigènes de HIV-2 et, pour certains d'entre eux également avec des antigènes de HIV-1 ou de ses variants, sont caractérisés en ce qu'ils ont également une structure peptidique en commun avec les antigènes de SIV. De façon avantageuse, ces peptides comprennent normalement au plus 40 résidus d'acides aminés.

Des peptides préférés sont les suivants :

env1

XRV-AIEKYL-DQA-LN-WGCAFQVCZ

env2

X-LE-AQI-QQEKNMYELOQLNZ

env3

XELGDYKLVEITPIG-APT--KR-----Z

env4

X----VTV-YGVP-WK-AT--LFCA-Z

env5

X---QE--L-NVTE-E--W-NZ

env6

XL---S-KPCVKLTPLCV--Z

env7

X---N-S-IT--C-K---Z

env8

X-I---YC-P-G-A-L-C-N-TZ

env9

X-----A-C-----W--Z

env10

35 X-G-DPE-----NC-GEF-YCN-----NZ

15

env11

X-----C-IKQ-I-----G---YZ

Plus particulièrement l'invention concerne les peptides suivants :

5 env1

XRV-AIEKYL-DQA-LN-WGCAFRQVCZ

env2

X-LE-AQIQQEKNMYELQKLNZ

env3

10 XELG DYKLVEITPIG-APT--KR----Z

env4

X----VTV-YGVP-W--AT--LFCA-Z

env5

X----E--L-NVTE-F--W-NZ

15 env6

XL---S-KPCVKL-PLC---Z

env7

X---N-S-I---C-K----Z

env8

20 X-I---YC-P-G-A-L-C-N-TZ

env9

X-----A-C-----W--Z

env10

X-G-DPE-----NC-GEF-YC-----NZ

25 env11

X-----C-I-Q-I-----G---YZ

Des peptides avantageux correspondant aux précédents, présentent les formules qui suivent :

env1

30 XRVTAIEKYLQDQARLNSWGCAFRQVCZ, ou

XRVTAIEKYLKDQAQLNAWGCAFRQVCZ

env2

XSLEQAQIQQEKNMYELQKLNSWZ, ou

XLLEEAQIQQEKNMYELQKLNSWZ

env3

XELG DYKLVEITPIGFAPTKEKRYSSAHZ, ou

XELG DYKLVEITPIGLAPTNVKRYTTG-Z

5 (On remarquera que les peptides env1, env2, env3 attestent de la très grande parenté entre HIV-2 et SIV-1. En effet le premier peptide est inclu dans le génome de HIV-2 et le second, dans celui de SIV-1).

env4

XabcdVTVeYGV PfWogAThiLFCAjZ,

10 dans lesquels les lettres de a à j peuvent avoir les significations suivantes :

a est C, E ou D

b est T, K, D, N ou I

c est Q ou L

d est Y ou W

15 e est F ou Y

f est T, V ou A

g est N ou E

h est I ou T

20 i est P ou T

j est T ou S

o est K ou R

env5

XabcoEdeLfNVTEgFhiWjNZ,

25 dans lequel les lettres de a à j peuvent avoir les significations suivantes :

a est D ou P

b est D ou N

c est Y ou P

30 d est I, V, I ou L

e est T, V, E ou A

f est V, G ou E ou -

g est A, N, G ou S

h est D ou N

35 i est A ou M

j est N, K ou E

o est Q ou S

env6

XLabcSdKPCVKLoPLCuefKZ,

5 dans lequel les lettres de a à f peuvent avoir les significations suivantes :

a est F ou W

b est E ou D

c est T ou Q

10 d est I ou L

e est A, S ou T .

f est M ou L

o est T ou S

u est V ou I

env7

XabCNxSyIocdCeKfghiz,

dans lequel les lettres de a à i et x et y peuvent avoir les significations suivantes :

a est N ou T ou I

20 b est H ou S ou N

c est E ou Q

d est S, A ou C

e est D ou P

f est H, V ou D

25 g est Y ou S

h est W ou F

i est D ou E

x est T ou R

y est V ou A

30 o est T ou Q

env8

XaIbcdYCxPeGfAgLhCiNjTZ,

dans lequel les lettres de a à k et x peuvent avoir les significations suivantes :

- a est A ou P
- b est R ou P
- c est F, I ou C
- d est R ou H
- e est P ou A
- 5 f est Y ou F
- g est L ou I
- h est R ou K
- i est - ou N
- j est D ou K
- 10 x est A ou T

env9

XwabcxyAdCefghizWjkZ,

15 dans lequel les lettres de a à k et x à z peuvent avoir les significations suivantes :

- a est K ou - ou E
- b est R ou -
- c est P ou M ou I
- d est W ou H ou Y
- 20 e est W ou N ou T ou R
- f est F ou I
- g est K ou S ou N ou G
- h est G ou R ou E
- i est - ou A ou T
- j est K ou N ou D ou S
- 25 k est D ou A ou N ou K ou E
- w est N, D ou I
- x est R ou G ou K
- y est Q ou K ou R
- z est K ou E ou Q ou N

env10

XaGbDPEcdefghNCiGEFjYCoKxlmnNZ,

30 dans lequel les lettres de a à n et x peuvent avoir les significations suivantes :

a est K ou - ou G
b est S ou G ou -
c est V ou I
d est A ou V ou T
5 e est Y ou T ou M ou F
f est M ou H
g est W ou S
h est T ou F
i est R ou G
10 j est L ou F
o est N ou K
k est M ou S
l est W ou Q ou K ou G
m est F ou L
15 n est L ou F
x est T ou S ou N

env11

XabcdwCeIoQfIxgyhizGjklyZ,

dans lequel les lettres de a à l et w à z peuvent avoir
20 les significations suivantes :

a est R ou T ou S ou N
b est N ou I
c est Y ou T
d est A ou L ou V
25 e est H ou R
f est I ou F
g est T ou M
h est H ou Q ou A
i est K ou E
j est R ou K
30 k est N ou A
l est V ou M
w est P ou Q
x est N ou K
35 y est W ou V

z est V ou T ou K

o est K ou R

La structure du peptide antigénique codé par le gène gag et désigné par gag1 est également représentée ci-après :

5 XDCKLVLKGLGaNPTLEEMLTAZ,

dans lequel la lettre a désigne M ou T.

Il sera remarqué que, d'une façon générale, les aminoacides ayant une signification univoque (donc représentés par une lettre majuscule correspondant à la nomenclature internationale) qui interviennent dans les définitions qui précèdent des peptides selon l'invention, se trouvent être la correspondance avec des aminoacides identiques placés dans le même ordre dans les séquences env ou gag correspondantes de la protéine env ou gag d'au moins l'un des HIV, ou de SIV-1.

Les positions de ces séquences sont soulignées et repérées au sein des séquences d'aminoacides des protéines env respectivement de HIV-2 ROD (CNCM n° I-532) et HIV-1 BRU (CNCM n° I-232) représentées à la figure 2. Par ailleurs, les alignements des acides aminés des protéines env et gag respectivement de SIV-1mac (CNCM n° I.521) et de HIV-2 ROD sont présentées à la figure 3 et à la figure 4.

Les traits pleins qui apparaissent en certaines localisations de ces séquences visent à souligner que certains aminoacides contenus dans ces séquences ont été volontairement déletés au plan de la présentation, afin de permettre la mise en alignement d'aminoacides respectivement identiques (alors marqués d'un astérisque) ou de deux points verticaux sur une même ligne verticale dans les séquences des protéines correspondantes de HIV-1 et de HIV-2 d'une part, de SIV et de HIV-2 d'autre part.

Outre les peptides précités, l'invention concerne également les peptides modifiés par insertion et/ou délétion et/ou substitution d'un ou plusieurs acides aminés, pour autant que les propriétés antigéniques ou immunogènes desdits peptides ne sont pas modifiées, ou que les propriétés de reconnaissance de l'antigène ou de l'anticorps avec lesdits peptides ne sont pas substantiellement modifiées.

Dans un mode de réalisation particulièrement préféré, l'invention concerne des peptides ayant des propriétés immunologiques en commun avec l'ossature peptidique de la glycoprotéine d'enveloppe des virus de la classe HIV-2, ces peptides contenant un nombre de résidus d'acides aminés n'excédant pas 40.

Ces peptides préférés selon l'invention ont les séquences suivantes :

env1

RVTAIEKYLQDQARLNSWGCAFRQVC

AIEKYLQDQ

RVSAIEKYLKDQAQLNAWGCAFRQVC

AIEKYLKDQ

env2

SLEQAAQIQQEKNMYELQKLNSW

QIQQEKN

LLEEAQIQQEKNMYELQKLNSW

env3

ELGDYKLVEITPIGFAPTKEKRYSSAH

YKLVEITPIGFAPTKEK

ELGDYKLVEITPIGLAPTNVKRYTTG-

YKLVEITPIGLAPTNVK

env4

CTQYVTVFYGVPTWKNATIPLFCAT

VTVFYGVPTWKNAT

CIQYVTVFYGVPAWRNATIPLFCAT

VTVFYGVPAWRNAT

EKLWVTVYYGVPVWKEATTLFCAS

VTVYYGVPVWKEAT

EDLWVTVYYGVPVWKEATTLFCAS

VTVYYGVPVWKEAT

5 DNLWVTVYYGVPVWKEATTLFCAS

VTVYYGVPVWKEAT

env5

DDYQEITL-NVTEAFDAWNN

L-NVTEAF

10 DDYSELAL-NVTESFDAWEN

L-NVTESF

PNPQEVLVNVTFNFNMWKN

LVNVTFNF

PNPQEIELENVTEGFNMWKN

LENVTEGF

15 PNPQEIALENVTFNFNMWKN

LENVTFNF

env6

ETSIKPCVKLTPLCVAMK

20 ETSIKPCVKLSPLCITMR

DQSLKPCVKLTPLCVSLK

DQSLKPCVKLTPLCVTLN

PCVKLTPLCV

env7

25 NHCNTSVITESCD

NTSVIT

NHCNTSVIQECCD

NTSVIQ

TSCNTSVITQACP

NTSVIT

30 INCNTSVITQACP

NTSVIT

INCNTSAITQACP

NTSAIT

env8

YCAPPGYALLRC-NDT

YCAPAGFAILKCNNKT.

YCAPAGFAILKCNDKK

5 YCAPAGFAILKCRDKK

env9

NKRPRQAWCWFKG-KWKD

NERPKQAWCRFGG-NWKE

N--MRQAHCNISRAKWNA

10 D--IRRAYCTINETEWDK

I--IGQAHCNISRAQWSK

env10

KGSDPEVAYMWTNCRGEFLYCNMTWFLN

NCRGEFLYCN

15 GG-DPEVTFMWTNCRGEFLYCKMNWFLN

NCRGEFLYCK

-GGDPEIVTHSFNCGGEFFYCNSTQLFN

NCGGEFFYCN

-GGDPEITTHSFNCRGGEFFYCNTSKLFN

NCRGEFFYCN

20 -GGDPEITTHSFNCGGEFFYNTSGLFN

NCGGEFFYCN

env11

RNYAPCHIKQIINTWHKVGRNVY

CHIKQII

25 RNYVPCHIRQIINTWHKVGVKNVY

CHIRQII

TITLPCRIKQFINMWQEVGKAMY

CRIKQFI

30 SITLPCRIKQIINMWQKTCKAMY

CRIKQII

NITLQCRIKQIIKMQAGR-KAIY

CRIKQII

gag1

35 DCKLVLKGLGTNPTELMLTA

Les peptides selon l'invention peuvent encore avantageusement être préparés par les techniques classiques, dans le domaine de la synthèse des peptides. Cette synthèse peut être réalisée en solution homogène ou en phase solide.

Par exemple, on aura recours à la technique de synthèse en solution homogène décrit par Houbenweyl dans l'ouvrage intitulé "Méthode der Organischen Chemie" (Méthode de la Chimie Organique) édité par E. Wunsch, vol. 15-I et II., THIEME, Stuttgart 1974.

Cette méthode de synthèse consiste à condenser successivement deux-à-deux les aminoacyles successifs dans l'ordre requis, ou à condenser des aminoacyles et des fragments préalablement formés et contenant déjà plusieurs aminoacyles dans l'ordre approprié, ou encore plusieurs fragments préalablement ainsi préparés, étant entendu que l'on aura eu soin de protéger au préalable toutes les fonctions réactives portées par ces aminoacyles ou fragments, à l'exception des fonctions amines de l'un et carboxyles de l'autre ou vice-versa, qui doivent normalement intervenir dans la formation des liaisons peptidiques, notamment après activation de la fonction carboxyle, selon les méthodes bien connues dans la synthèse des peptides. En variante, on pourra avoir recours à des réactions de couplage mettant en jeu des réactifs de couplage classique, du type carbodiimide, tels que par exemple la 1-éthyl-3-(3-diméthyl-amino-propyl)-carbodiimide. Lorsque l'aminoacyle mis en oeuvre possède une fonction acide supplémentaire (notamment dans le cas de l'acide glutamique), ces fonctions seront par exemple protégées, par des groupes t-bustylester.

Dans le cas de la synthèse progressive, acide aminé par acide aminé, la synthèse débute de préférence par la condensation de l'amino-acide C-terminal avec l'aminoacide qui correspond à l'aminoacyle voisin dans

la séquence désirée et ainsi de suite, de proche en proche, jusqu'à l'acide aminé N-terminal. Selon une autre technique préférée de l'invention, on a recours à celle décrite par R.D. MERRIFIELD dans l'article intitulé "Solid phase peptide synthesis" (J. Am. Soc., 5 45, 2149-2154).

Pour fabriquer une chaîne peptidique selon le procédé de MERRIFIELD, on a recours à une résine polymère très poreuse, sur laquelle on fixe le premier acide aminé C-terminal de la chaîne. Cet acide aminé est fixé 10 sur la résine par l'intermédiaire de son groupe carboxylique et sa fonction amine est protégée, par exemple par le groupe t-butyloxycarbonyle.

Lorsque le premier acide aminé C-terminal est 15 ainsi fixé sur la résine, on enlève le groupe protecteur de la fonction amine en lavant la résine avec un acide.

Dans le cas où le groupe protecteur de la fonction amine est le groupe t-butyloxycarbonyle, il peut être éliminé par traitement de la résine à l'aide 20 d'acide trifluoroacétique.

On couple ensuite le deuxième acide aminé qui fournit le second amino-acyle de la séquence recherché, à partir du résidu amino-acyle C-terminal sur la fonction amine déprotégée du premier acide aminé C-terminal fixé sur la chaîne. De préférence, la fonction carboxyle 25 de ce deuxième acide aminé est activée, par exemple par la dicyclohexylcarbodiimide, et la fonction amine est protégée, par exemple par le t-butyloxycarbonyle.

On obtient ainsi la première partie de la 30 chaîne peptidique recherchée, qui comporte deux acide aminés, et dont la fonction amine terminale est protégée. Comme précédemment, on déprotège la fonction amine et on peut ensuite procéder à la fixation du troisième aminoacyle, dans les conditions analogues à celles de l'addition du deuxième acide aminé C-terminal. 35

On fixe ainsi, les uns après les autres, les acides aminés qui vont constituer la chaîne peptidique sur le groupe amine chaque fois déprotégé au préalable de la portion de la chaîne peptidique déjà formée, et qui est rattachée à la résine.

5

Lorsque la totalité de la chaîne peptidique désirée est formée, on élimine les groupes protecteurs des différents acide aminés constituant la chaîne peptidique et on détache le peptide de la résine par exemple à l'aide d'acide fluorydrique.

10

L'invention concerne également les oligomères hydrosolubles des peptides monomères sus-indiqués. L'oligomérisation peut provoquer un accroissement de l'immunogénicité des peptides monomères selon l'invention. Sans qu'une telle indication chiffrée puisse être considérée comme limitative, on mentionnera néanmoins que ces oligomères peuvent, par exemple, contenir de 2 à 10 unités monomères.

15

Les unités monomères entrant dans cet oligomère sont soit toutes constituées par le polypeptide de séquence 1 ou par le polypeptide de séquence 2, soit par l'un et l'autre de ces polypeptides.

20

On peut avoir recours, pour réaliser l'oligomérisation, à toute technique de polymérisation couramment utilisée dans le domaine des peptides, cette polymérisation étant conduite jusqu'à l'obtention d'un oligomère ou polymère contenant le nombre de motifs monomères requis pour l'acquisition de l'immunogénicité désirée.

25

30

Une méthode d'oligomérisation ou de polymérisation du monomère consiste dans la réaction de celui-ci avec un agent de réticulation tel que le glutaraldéhyde.

35

..

On peut également avoir recours à d'autres méthodes d'oligomérisation ou de couplage, par exemple à

celle mettant en jeu des couplages successifs d'unités monomères, par l'intermédiaire de leurs fonctions terminales carboxyle et amine en présence d'agents de couplage homo- ou hétéro-bifonctionnels.

On peut également pour la production de molécules comportant un ou plusieurs motifs de 17 acides aminés tels que définis ci-dessus, avoir recours à des techniques du génie génétique mettant en oeuvre des micro-organismes transformés par un acide nucléique déterminé comprenant des séquences nucléotidiques appropriées correspondantes.

L'invention concerne également les acides nucléiques contenant une ou plusieurs séquences issues de la séquence de l'ADNc du virus HIV-2 ROD. Ces séquences repérées par la numérotation figurant sur la séquence précédemment décrite, codent pour certains peptides intéressants de l'invention.

Séquence codant pour env1 nucléotides 7850 à 7927

| | | | | | |
|----|---|---|--------------|---|-------------|
| | " | " | <u>env2</u> | " | 8030 à 8095 |
| 20 | " | " | <u>env3</u> | " | 7601 à 7636 |
| | " | " | <u>env4</u> | " | 6170 à 6247 |
| | " | " | <u>env5</u> | " | 6294 à 6349 |
| | " | " | <u>env6</u> | " | 6392 à 6445 |
| | " | " | <u>env7</u> | " | 6724 à 6763 |
| 25 | " | " | <u>env8</u> | " | 6794 à 6838 |
| | " | " | <u>env9</u> | " | 7112 à 7162 |
| | " | " | <u>env10</u> | " | 7253 à 7336 |
| | " | " | <u>env11</u> | " | 7358 à 7426 |
| | " | " | <u>gag1</u> | " | 1535 à 1597 |

30 L'invention concerne enfin les acides nucléiques correspondants du virus SIV, contenant une ou plusieurs séquences issues de l'ADNc du virus SIV-1. Ces séquences codant pour les peptides env1 à env11 et gag1 peuvent être repérés sur la figure 3 par comparaison 35 avec les séquences correspondantes décrites pour HIV-2.

Il va de soi que l'invention concerne les acides nucléiques correspondant à des séquences placées en des régions analogues des ADNC dérivés de variants de HIV-2 ROD ou de SIV, ainsi que tous les acides nucléiques dont les modifications vis à vis des précédents résulteraient de la mise à profit de la dégénérescence du code génétique.

L'invention concerne encore les conjugués obtenus par couplage covalent des peptides selon l'invention (ou des susdits oligomères) à des molécules porteuses (naturelles ou synthétiques), physiologiquement acceptables et non toxiques, par l'intermédiaire de groupements réactifs complémentaires respectivement portés par la molécule porteuse et le peptide. Des exemples de groupements appropriés sont illustrés dans ce qui suit :

A titre d'exemple de molécules porteuses ou supports macromoléculaires entrant dans la constitution des conjugués selon l'invention, on mentionnera des protéines naturelles, telles que l'anatoxine tétanique, l'ovalbulmine, des sérums albumines, des hémocytamines, etc...

A titre de support macromoléculaires synthétiques, on mentionnera par exemple des polylysines ou des poly(D-L-alanine)-poly(L-lysine).

La littérature mentionne d'autres types de supports macromoléculaires susceptibles d'être utilisés, lesquels présentent en général un poids moléculaire supérieur à 20 000.

Pour synthétiser les conjugués selon l'invention, on peut avoir recours à des procédés connus en soi, tels que celui décrit par FRANTZ et ROBERTSON dans Infect. and Immunity, 33, 193-198 (1981), ou celui décrit dans Applied and Environmental Microbiology, (octobre 1981), vol. 42, n° 4, 611-614 par P.E. KAUFFMAN

en utilisant le peptide et la molécule porteuse appropriée.

Dans la pratique, on utilisera avantageusement comme agent de couplage les composés suivants, cités à titre non limitatif : aldéhyde glutarique, chloroformiate d'éthyle, carbodiimides hydrosolubles [N-éthyl-N'(3-diméthylamino-propyl) carbodiimide, HCl], diisocyanates, bis-diazobenzidine, di- et trichloro-s-triazines, bromures de cyanogène, ainsi que les agents de couplage mentionnés dans Scand. J. Immunol., (1978), vol. 8, p. 7-23 (AVRAMEAS, TERNYNCK, GUESDON).

On peut avoir recours à tout procédé de couplage faisant intervenir d'une part une ou plusieurs fonctions réactives du peptide et d'autre part, une ou plusieurs fonctions réactives de molécules supports. Avantageusement, il s'agit des fonctions carboxyle et amine, lesquelles peuvent donner lieu à une réaction de couplage en présence d'un agent de couplage du genre de ceux utilisés dans la synthèse des protéines, par exemple, le 1-éthyl-3-(3-diméthylaminopropyl)-carbodiimide, le N-hydroxybenzotriazole, etc... On peut encore avoir recours à la glutaraldéhyde, notamment lorsqu'il s'agit de relier entre eux des groupes aminés respectivement portés par le peptide et la molécule support.

Les peptides selon l'invention possèdent des propriétés antigéniques. Ils peuvent donc être utilisés dans des procédés de diagnostic pour la détection d'une infection par le virus HIV-2.

Comme on l'a déjà mentionné, des études ont permis de distinguer deux groupes de peptides pouvant être mis en oeuvre dans des procédés de détection d'anticorps contre le virus HIV-2 dans un fluide biologique humain, notamment un sérum ou un liquide céphalo-rachidien.

Un premier groupe (I) comprend les peptides gag₁. Ces peptides reconnaissent des anticorps anti-HIV-2 et sont donc capables de détecter une infection par HIV-2. Ils reconnaissent également dans une certaine mesure des anticorps anti-HIV-1.

5 Un second groupe (II) comprend des peptides qui correspondent plus particulièrement à ceux qui sont situés dans la partie transmembranaire et dans la fin de la partie externe de la protéine d'enveloppe. Ces peptides sont ceux précédemment désignés par env1, env2 et env3. Ils permettent la reconnaissance spécifique de la présence d'anticorps contre HIV-2 et permettent donc de discriminer chez une personne les infections passées ou présentes dues à un HIV, plus particulièrement entre 10 celles qui ont été provoquées par un HIV-2 et celles qui 15 l'ont été par un HIV-1.

L'invention concerne également une composition contenant au moins l'un des susdits peptides ou au moins 20 un oligomère de ce peptide, caractérisée en ce qu'elle a la capacité d'être reconnue par des sérum s d'origine humaine contenant des anticorps contre le virus HIV-2.

L'invention concerne un procédé de diagnostic in vitro un ou des peptides selon l'invention pour la détection d'anticorps contre HIV-2 dans des fluides biologiques, en particulier dans des sérum s humains.

25 D'une façon générale le procédé de diagnostic in vitro ci-dessus comprend les étapes suivantes :

- la mise en contact de ce liquide biologique avec lesdits peptides;
- la détection de la présence éventuelle d'un complexe peptidé-anticorps par des méthodes physiques ou chimiques, dans ledit liquide biologique.

Dans un mode de réalisation préféré de l'invention, la détection du complexe antigène-anticorps est 30 réalisée grâce à des tests immunoenzymatiques (du type

ELISA), immunofluorescents (du type IFA), radioimmunologiques (du type RIA) ou des tests de radioimmunoprécipitation (du type RIPA).

Ainsi l'invention concerne également tout peptide selon 5 l'invention marqué à l'aide d'un marqueur adéquat du type enzymatique, fluorescent, radioactif, etc...

De telles méthodes comprennent par exemple les étapes suivantes :

- dépôt de quantités déterminées d'une 10 composition peptidique selon l'invention dans les puits d'une microplaque de titration,

- introduction dans lesdits puits de dilutions croissantes du sérum devant être diagnostiqué,

- incubation de la microplaque,

- rinçages répétés de la microplaque,

- 15 introduction dans les puits de la microplaque d'anticorps marqués contre des immunoglobulines du sang, le marquage de ces anticorps ayant été réalisé à l'aide d'une enzyme sélectionnée parmi celles qui sont 20 capables d'hydrolyser un substrat en modifiant l'absorption des radiations de ce dernier, au moins à une longueur d'onde déterminée,

- détection, en comparaison avec un témoin de contrôle, de la quantité de substrat hydrolysé.

25 L'invention concerne également des coffrets ou kits pour le diagnostic in vitro de la présence d'anticorps contre les virus HIV-2 et, dans certains cas, HIV-1 dans un milieu biologique qui comprennent :

- une composition peptidique selon l'invention, 30

- les réactifs pour la constitution du milieu propice à la réalisation de la réaction immunologique,

- les réactifs permettant la détection du complexe antigènes-anticorps produit par la réaction immunologique. De tels réactifs peuvent également porter 35

un marqueur, ou être susceptibles d'être reconnus à leur tour par un réactif marqué. Plus particulièrement dans le cas où la composition polypeptidique sus-mentionnée n'est pas marquée.

5 - un tissu fluide biologique de référence dépourvu d'anticorps reconnus par la composition polypeptidique sus-mentionnée,

L'invention concerne les anticorps eux-mêmes formés contre les peptides de l'invention.

10 Il va de soi que cette production n'est pas limitée aux anticorps polyclonaux.

15 Elle s'applique encore à tout anticorps monoclonal produit par tout hybridome susceptible d'être formé, par des méthodes classiques, à partir des cellules spléniques d'un animal, notamment de souris ou de rat, immunisés contre l'un des peptides de l'invention, d'une part et des cellules d'une lignée de cellule myélome approprié d'autre part, et d'être sélectionné, 20 par sa capacité à produire des anticorps monoclonaux reconnaissant le peptide initialement mis en oeuvre pour l'immunisation des animaux.

25 L'invention concerne également des compositions immunogènes pour la production de vaccins dont le principe actif est constitué par au moins un peptide selon l'invention, ou un oligomère de ce peptide, ou un peptide sous forme conjuguée avec une molécule porteuse, caractérisées en ce qu'elles induisent la production 30 d'anticorps contre les susdits peptides en quantité suffisante pour aussi inhiber les protéines du rétrovirus HIV-2, voire même le rétrovirus HIV-2 entrant en association avec un véhicule pharmaceutiquement acceptable.

35 Les compositions immunogènes pour la production de vaccins comprennent de façon avantageuse plus particulièrement au moins l'un des peptides précédemment désignés par env4, env5, env6, env7, env8, env9, env10,

env11 voir des mélanges de ceux-ci.

Parmi ces peptides aptes à constituer des principes actifs de vaccins certains sont particulièrement préférés car ils possèdent une structure de base en acides aminés correspondant à des régions des glycoprotéines d'enveloppe qui présentent un important degré de conservation, non seulement dans les HIV-2, et dans les SIV, mais également dans les HIV-1. Ces peptides particulièrement préférés sont les peptides désignés par 10 env4, certains peptides env5, env6 et env10.

Dans un mode de réalisation préféré de l'invention les peptides immunogènes (ou fragments de ces peptides) aptes à constituer des principes actifs de vaccins sont choisis parmi ceux dont les formules correspondent à des séquences qui, dans les glycoprotéines d'enveloppe de HIV-2, SIV et HIV-1 présentant une homologie en acides aminés supérieure à 50%, qui appartiennent à la partie externe de l'enveloppe du virus, qui sont dépourvus ou presque de délétions, et qui renferment des résidus de cystéine favorables à la stabilisation des liaisons et à la constitution de boucles d'ancre.

Les peptides suivants appartiennent à cette catégorie de peptides préférés.

25 env4

XVTV-YGVP-W--ATZ

env5

XL-NVTE-FZ

env6

30 XKPCVKL-PLC-Z

env7

XN-S-I-Z

env10

XNC-GEF-YC-Z

env11

XC-I-Q-I2

Des compositions pharmaceutiques avantageuses sont constituées par des solutions, suspensions ou liposomes injectables contenant une dose efficace d'au moins un produit selon l'invention. De préférence, ces solutions, suspensions ou liposomes sont réalisés dans une phase aqueuse stérilisée isotonique, de préférence saline ou glucosée.

L'invention concerne plus particulièrement de telles suspensions, solutions ou forme liposome qui sont aptes à être administrées par injections intradermiques, intramusculaires ou sous-cutanées, ou encore par scarifications.

Elle concerne également des compositions pharmaceutiques administrables par d'autres voies, notamment par voie orale.

Les compositions pharmaceutiques selon l'invention, utilisables en tant que vaccins pour être efficaces dans la production d'anticorps contre le virus HIV-2, peuvent à titre d'exemple être administrées à des doses situées entre 10 et 500 µg/kg, de peptides selon l'invention, de préférence de 50 à 100 µg/kg.

Ces doses sont citées à titre d'exemple et ne possèdent en aucun cas un caractère limitatif.

Comme on l'a déjà indiqué plus haut les différents peptides qui ont été définis peuvent comprendre des modifications qui n'ont pas pour effet de modifier de façon fondamentale leurs propriétés immunologiques. Les peptides équivalents qui en résultent entrent dans le champ des revendications qui suivent. A titre d'exemples de peptides équivalents on mentionnera ceux dont les structures en correspondance avec des régions des ADNc d'autres variants de HIV-2 de SIV ou de HIV-1, lorsque ces régions ont été mises en alignement dans des

35

conditions semblables à celles qui ont été évoquées ci-dessus, à propos de HIV-2 ROD, SIV et HIV-1 BRU. A titre d'autres de ces peptides, on mentionnera ceux dont les structures sont en correspondance avec de telles régions dans les ADNC qui ont fait l'objet de dépôts à la CNCM, notamment sous les numéros I-502, I-642 (HIV-2 IRMO), I-643 (HIV-2 EHO) ainsi que, dans les cas appropriés, des variants de HIV-1 qui ont fait l'objet de dépôts à la CNCM sous les numéros I-232, I-240, I-241, I-550, I-551.

Les peptides selon l'invention peuvent encore être définis par les formules suivantes (dans lesquels X, Z et les tirets "-" ont les significations sus-indiquées) :

15

20

25

30

35

36

XRV-AIEKYL-DQA-LN-WGCAFQVCZ
XAIEKYL-D2

X-LE-AQIQQEKNMYELQKLNSWZ
5 XQIQQEKNZ

XELG DYKLVEITPIG-APT--KR----Z
XYKLVEITPIG-APT--KRZ

10 X----VTV-YGVP-W--AT--LFCA-Z
XVTV-YGVP-W--ATZ

X----E--L-NVTE-F--W-NZ
XL-NVTE-FZ

15 XL---S-KPCVKL-PLC----Z
XKPCVKL-PLC-Z
XS-KPCVKL-PLC-Z

20 X---N-S-I---C-Z
XN-S-I-Z

XYC-P-G-A-L-C-N-TZ

25 X-----A-C-----W--Z

NKRPRQAWCWFKG-KWKD

X-G-DPE-----NC-GEF-YC-----NZ

30 X-----C-I-Q-I-----G---YZ

• L'invention concerne également outre les peptides de SIV déjà décrits, les protéines codées par l'ADNc du virus SIV. Elle concerne également les protéines de tout virus immunologiquement étroitement apparenté à SIV-1mac, en particulier tout virus dont les protéines et les glycoprotéines d'enveloppe croisent immunologiquement et dont les ADNc présentent un pourcentage d'homologie d'au moins 95% et de préférence d'au moins 98%.

En particulier l'invention concerne :

- 10 1/ les protéines et glycoprotéines de l'enveloppe codées par le gène env et représentées à la figure 3,
- 2/ la protéine GAG représentée à la figure 4,
- 15 3/ la protéine POL représentée à la figure 5,
- 4/ la protéine Q représentée à la figure 6,
- 20 5/ la protéine R représentée à la figure 7,
- 6/ la protéine X représentée à la figure 8,
- 7/ la protéine F représentée à la figure 9,
- 25 8/ la protéine TAT représentée à la figure 10,

Les acides aminés des protéines précitées de SIV, ont été représentées en alignement avec les séquences d'acides aminés des protéines correspondantes du virus HIV-2 ; les points verticaux figurant entre les deux séquences correspondent aux acides aminés communs entre les protéines des deux virus.

Les séquences d'ADNc codant pour les protéines précitées apparaissent sur la figure 1B. L'invention concerne, outre les séquences nucléiques précitées toute séquence nucléique modifiée, qui code également pour les protéines du rétrovirus SIV ou d'un variant.

Ces séquences d'ADNc repérées par la numérotation figurant sur les séquences décrites précédemment (figure 1B) sont les suivantes :

L'invention concerne donc naturellement les protéines précédemment décrites, lorsqu'elles sont obtenues à partir du virus SIV ou lorsqu'elles sont préparées par une méthode de synthèse, notamment par l'une des méthodes déjà citées en rapport avec la synthèse des peptides de plus petite taille.

L'invention concerne également l'utilisation des protéines précédentes pour le diagnostic de la présence éventuelle d'anticorps dirigés contre les protéines de HIV-2, voire contre HIV-2 en entier, ou pour certaines d'entre elles l'utilisation aux fins de diagnostic d'une infection due à l'un des virus HIV. Ainsi le peptide GAG codé par le gène correspondant peut être utilisé pour repérer la présence éventuelle d'anticorps anti-HIV-1 ou anti-HIV-2. Les protéines ENV sont utilisées de préférence pour le diagnostic spécifique d'une infection due à HIV-2 ou un de ses variants, parfois pour le diagnostic d'une infection par HIV-2 ou HIV-1.

L'invention concerne donc également un procédé
de diagnostic in vitro de détection d'anticorps contre
HIV-2 et éventuellement contre HIV-1 dans des fluides-
biologiques et en particulier dans des sérums humains.
De tels procédés applicables pour l'utilisation des pro-
téines précédentes de SIV comme protéines de diagnostic,

ont déjà été décrits dans la présente invention.

L'invention concerne aussi des coffrets ou "kits" pour le diagnostic in vitro de la présence d'anticorps le virus HIV-2 et dans certains cas contre HIV-1 dans un milieu biologique. De tels kits mettant en oeuvre les peptides précédents ont également été décrits dans la présente invention.

L'invention concerne également des compositions immunogènes pour la production de vaccins, dont le principe actif est constitué de façon avantageuse par au moins la partie de la protéine ENV du virus SIV, cette protéine pouvant être sous forme conjuguée avec une molécule porteuse. Ces compositions immunogènes induisent la production d'anticorps contre le susdit peptide en quantité suffisante pour inhiber les protéines du rétrovirus HIV-2, voire le rétrovirus HIV-2 lui-même.

Toutefois l'utilisation aux fins de diagnostic des protéines de SIV n'est en rien limitée à celle des seuls protéines ENV ou GAG. D'autres protéines parmi celles décrites peuvent être envisagées, pour préparer des compositions de diagnostic voire de vaccin.

25

30

35

REVENDICATIONS

1/ Peptide ayant des propriétés immunologiques en commun avec celles de l'ossature peptidique de la glycoprotéine d'enveloppe des virus de la classe HIV-2, caractérisé en ce qu'il a également une structure peptidique en commun avec l'ossature peptidique de la glycoprotéine de SIV.1.

2/ Peptide ayant des propriétés immunologiques en commun avec celles de l'ossature peptidique de la glycoprotéine d'enveloppe des virus de la classe HIV-2, ces peptides contenant un nombre de résidus d'acides aminés n'excédant pas 40, caractérisé en ce qu'il a également une structure peptidique en commun avec l'ossature peptidique de la glycoprotéine de SIV.1.

3/ Peptide selon la revendication 2 caractérisé par l'une des formules :

XRV-AIEKYL-DQA-LN-WGCAFRQVCZ

XAIKEYL-DZ

dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

RVTAIEKYLQDQARLNSWGCAFRQVC

AIEKYLQDQ

RVSAIEKYLKDQAQQLNAWGCAFRQVC

AIEKYLKDQ

4/ Peptide selon la revendication 2 caractérisé par l'une des formules :

X-LE-AQIQQEKNMYELQKLNSWZ

XQIQQEKNZ

35 dans laquelle X et Z sont des groupements OH ou NH₂ ou,

dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

SLEQAAQIQQEKNMYELQKLNSW

QIQQEKN

LLEEAQIQQEKNMYELQKLNSW

5 / Peptide selon la revendication 2 caractérisé par l'une des formules :

XELGDYKLVEITPIG-APT--KR-----Z

XYKLVEITPIG-APT--KRZ

15 dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

ELGDYKLVEITPIGFAPTKEKRYSSAH

YKLVEITPIGFAPTKEK

ELGDYKLVEITPIGLAPTNVKRYTTG-

YKLVEITPIGLAPTNVK

6 / Peptide selon la revendication 2 caractérisé par l'une des formules :

X----VTV-YGVP-W--AT--LFCA-Z

XVTW-YGVP-W--ATZ

30 dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond

42

à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

CTQYVTVFYGVPTWKNATIPLFCAT

5 VTVFYGVPTWKNAT

CIQYVTVFYGVPAWRNATIPLFCAT

VTVFYGVPAWRNAT

EKLWVTVYYGVPVVKEATTLFCAS

VTVYYGVPVVKEAT

10 7/ Peptide selon la revendication 6 caractérisé par l'une des formules :

CTQYVTVFYGVPTWKNATIPLFCAT

VTVFYGVPTWKNAT

CIQYVTVFYGVPAWRNATIPLFCAT

VTVFYGVPAWRNAT

EKLWVTVYYGVPVVKEATTLFCAS

VTVYYGVPVVKEAT

EDLWVTVYYGVPVVKEATTLFCAS

VTVYYGVPVVKEAT

20 DNLWVTVYYGVPVVKEATTLFCAS

VTVYYGVPVVKEAT

8/ Peptide selon la revendication 2 caractérisé par l'une des formules :

X---E--L-NVTE-F--W-NZ

25 XL-NVTE-FZ

dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

DDYQEITL-NVTEAFDAWNN

35 L-NVTE

DDYSELAL-NVTESFDAWEN
PNPQEVLVNVTENFNMWKN
LVNVTE

9/ Peptide selon la revendication 8 caractérisé
5 par l'une des formules :

DDYQEITL-NVTEAFDAWNN
L-NVTEAF

DDYSELAL-NVTESFDAWEN
L-NVTESF

10 PNPQEVLVNVTENFNMWKN
LVNVTENF

PNPQEIELENVTEGFNMWKN
LENVTEGF

PNPQEIALENVTENFNMWKN
LENVTENF

15 10/ Peptide selon la revendication 2 caractérisé
par l'une des formules :

XL---S-KPCVKL-PLC---Z
XKPCVKLTPLCVZ

20 20 dans laquelle X et Z sont des groupements OH ou NH₂ ou,
dans la mesure où les propriétés immunologiques du pep-
tide dépourvu de ces groupes ne s'en trouvent pas essen-
tiellement modifiées, des groupes comportant de 1 à 5
25 résidus d'acides aminés, et chacun des tirets correspond
à un résidu aminoacyle choisi parmi ceux qui permettent
de conserver au peptide sus-défini les propriétés immu-
nologiques de l'une des séquences suivantes :

LFETSIKPCVKLTPLCVAMK

30 LFETSIKPCVKLSPLCITMR
LWDQSLKPCVKLTPLCVSLK

KPCVKLTPLCV

KPCVKLSPLCI

SLKPCVKLTPLCV

11/ Peptide selon la revendication 10 caractérisé par l'une des structures suivantes :

LEETSIKPCVKLTPLCVAMK

LFETSIKPCVKLSPLCITMR

5 LWDQSLKPCVKLTPLCVSLK

LWDQSLKPCVKLTPLCVTLN

PCVKLTPLCV

KPCVKLSPLCI

12/ Peptide selon la revendication 2 caractérisé en ce qu'il contient la structure de base :

10 X---N-S-I---C-Z

XN-S-I-Z

15 dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

20 NHCNTSVITESCD

NTSVIT

NHCNTSVIQECCD

NTSVIQ

25 TSCNTSVITQACP

NTSVIT

13/ Peptide selon la revendication 12 caractérisé par l'une des formules suivantes :

NHCNTSVITESCD

30 NTSVIT

NHCNTSVIQECCD

NTSVIQ

TSCNTSVITQACP

NTSVIT

35 INCNTSVITQACP

NTSVIT

INCNTSAITQACP

NTSAIT

14/ Peptide selon la revendication 2 caractérisé
5 par l'une des formules suivantes :

XYC-P-G-A-L-C-N-TZ

dans laquelle X et Z sont des groupements OH ou NH₂ ou,
dans la mesure où les propriétés immunologiques du
peptide dépourvu de ces groupes ne s'en trouvent pas es-
10 sentiellement modifiées, des groupes comportant de 1 à 5
résidus d'acides aminés, et chacun des tirets correspond
à un résidu aminoacyle choisi parmi ceux qui permettent
de conserver au peptide sus-défini les propriétés immu-
nologiques de l'une des séquences suivantes :

15 YCAPPGYALLRC-NDT

YCAPAGFAILKCNNKT

15/ Peptide selon la revendication 14 caractérisé
par l'une des formules :

YCAPPGYALLRC-NDT

20 YCAPAGFAILKCNNKT

YCAPAGFAILKCNDDKK

YCAPAGFAILKCRDKKK

16/ Peptide selon la revendication 2 caractérisé
par la formule :

25 X-----A-C-----W--Z

dans laquelle X et Z sont des groupements OH ou NH₂ ou,
dans la mesure où les propriétés immunologiques du
peptide dépourvu de ces groupes ne s'en trouvent pas es-
sentiellement modifiées, des groupes comportant de 1 à 5
30 résidus d'acides aminés, et chacun des tirets correspond
à un résidu aminoacyle choisi parmi ceux qui permettent
de conserver au peptide sus-défini les propriétés immu-
nologiques de l'une des séquences suivantes :

NKRPRQAWCWFKG-KWKD

35 NERPKQAWCRFCGG-NWKE

N--MRQAHCNISRAKWNA

17/ Peptide selon la revendication 16 caractérisé par la formule suivante :

NKRPRQAWCWFKG-KWKD

5 NERPKQAWCRFGG-KWKE

N--MRQAHCNISRAKWNA

D--IRRAYCTINETEWDK

I--IGQAHCNISRAQWSK

18/ Peptide selon la revendication 2 caractérisé 10 par la formule suivantes :

X-G-DPE-----NC-GEF-YC-----N₂

XNC-GEF-YC-Z

dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du 15 peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

20 KGSDPEVAYMWTNCRGEFLYCNMTWFLN

NCRGEFLYCN

GG-DPEVTFMWTNCRGEFLYCKMNWFLN

NCRGEFLYCK

25 -GGDPEIVTHSFNCGGEFFYCNSTQLFN

NCGGEFFYCN

19/ Peptide selon la revendication 18 caractérisé par l'une des structures suivantes :

KGSDPEVAYMWTNCRGEFLYCNMTWFLN

NCRGEFLYCN

30 GG-DPEVTFMWTNCRGEFLYCKMNWFLN

NCRGEFLYCK

-GGDPEIVTHSFNCGGEFFYCNSTQLFN

NCGGEFFYCN

35 -GGDPEITTHSFNCRGEFFYCNTSKLFN

NCRGEFFYCN

-GGDPEITTHSFNCGGEFFYCNTSGLFN

NCGGEFFYCN

20/ Peptide selon la revendication 2 caractérisé par l'une des formules suivantes :

5 X-----C-I-Q-I-----G---YZ

XC-I-Q-IZ

dans laquelle X et Z sont des groupements OH ou NH₂ ou, dans la mesure où les propriétés immunologiques du peptide dépourvu de ces groupes ne s'en trouvent pas essentiellement modifiées, des groupes comportant de 1 à 5 résidus d'acides aminés, et chacun des tirets correspond à un résidu aminoacyle choisi parmi ceux qui permettent de conserver au peptide sus-défini les propriétés immunologiques de l'une des séquences suivantes :

15 RNYAPCHIKQIINTWHKVGRNVY

CHIKQII

RNYVPCHIRQIINTWHKVGRNVY

CHIRQII

20 TITLPCRIKQFINMWQEVGKAMY

CRIKQFI

21/ Peptide selon la revendication 20 caractérisé par l'une des structures suivantes :

RNYAPCHIKQIINTWHKVGRNVY

CHIKQII

25 RNYVPCHIRQIINTWHKVGRNVY

CHIRQII

TITLPCRIKQFINMWQEVGKAMY

CRIKQFI

30 SITLPCRIKQIINMWQKTCKAMY

CRIKQII

NITLQCRIKQIIKMVAGR-KAIY

CRIKQII

35 22/ Peptide antigénique gag1, caractérisé par l'une des structures de base :

XDCKLVLKGLGMNPTLEEMLTAZ

XDCKLVLKGLGTNPTLEEMLTAZ

dans lesquelles X et Z sont des groupements OH ou NH₂
ou, dans la mesure où les propriétés immunologiques du
5 peptide dépourvu de ces groupes ne s'en trouvent pas es-
sentiellement modifiées, des groupes comportant de 1 à 5
résidus d'acides aminés, et dans lesquelles chacun des
tirets correspond à un résidu aminoacyle choisi parmi
ceux qui permettent de conserver au peptide sus-défini
10 les propriétés immunologiques de l'une ou l'autre des
séquences suivantes :

DCKLVLKGLGMNPTLEEMLTA

DCKLVLKGLGTNPTLEEMLTA

23/ Séquence nucléotidique caractérisée en ce
15. qu'elle renferme tout ou partie de la séquence d'acides
nucléiques définie à la figure 1B.

24/ Séquence nucléotidique caractérisée en ce
qu'elle renferme tout ou partie de la séquence d'acides
nucléiques définie à la figure 1C.

20 25/ Séquence nucléotidique selon la revendication
23, caractérisée en ce qu'elle comprend les séquences
nucléotidiques :

GAG s'étendant entre les nucléotides 550 à 2068

| | | | | | |
|----|-------|---|---|---|-----------------------|
| | POL | " | " | " | 1726 à 4893 |
| 25 | Q | " | " | " | 4826 à 5467 |
| | X | " | " | " | 5298 à 5633 |
| | R | " | " | " | 5637 à 5939 |
| | F | " | " | " | 8569 à 9354 |
| | TAT-1 | " | " | " | 5788 à 6084 |
| 30 | ART-1 | " | " | " | 6014 à 6130 |
| | TAT-2 | " | " | " | 8296 à 8391 |
| | ART-2 | " | " | " | 8294 à 8548 |
| | LTR | " | " | " | 8950 à 9468 <u>et</u> |
| | | | | | 1 à 316 |
| 35 | ENV | " | " | " | 6090 à 8732 |

26/ Peptide ayant une structure peptidique en commun avec l'ossature peptidique de SIV-1, caractérisé en ce qu'il comprend tout ou partie des séquences d'acides aminés parmi les séquences suivantes :

5 ENV représentée à la figure 3

| | | | |
|------------|---|---|----|
| <u>GAG</u> | " | " | 4 |
| <u>POL</u> | " | " | 5 |
| Q | " | " | 6 |
| R | " | " | 7 |
| X | " | " | 8 |
| F | " | " | 9 |
| TAT | " | " | 10 |
| ART | " | " | 11 |

10 27/ Acide nucléique recombinant caractérisé en ce qu'il comprend la totalité ou une partie d'un ADNc selon 15 une quelconque des revendications 23 à 25, inséré dans un acide nucléique provenant d'un vecteur.

28/ Acide nucléique recombinant selon la revendication 27, caractérisé en ce qu'il est marqué.

20 29/ Composition antigénique contenant le peptide gag selon la revendication 26 ou 27, au moins un peptide gag1 selon la revendication 22 ou au moins un oligomère de ce peptide, caractérisée en ce qu'elle a la capacité 25 d'être reconnue par des fluides biologiques d'origine humaine, notamment des sérums contenant des anticorps anti-HIV-2 et dans une certaine mesure des anticorps anti-HIV-1.

30 30/ Composition antigénique contenant le peptide env selon la revendication 26 ou au moins un peptide selon les revendications 3, 4 et 5 ou au moins un oligomère de ce peptide, caractérisée en ce qu'elle reconnaissent spécifiquement la présence d'anticorps contre HIV-2.

31/ 35/ Composition immunogène contenant tout ou partie du peptide env selon la revendication 26 ou au moins

un peptide ou au moins un oligomère de ce peptide ou ce peptide sous forme conjuguée avec une molécule porteuse, selon les revendications 6 à 21, en association avec un véhicule pharmaceutique acceptable pour la production de vaccins, caractérisée en ce qu'elle induit la production d'anticorps contre les susdits peptides en quantité suffisante pour inhiber efficacement les protéines du rétrovirus HIV-2, voire même le rétrovirus HIV-2 entier.

32/ Composition immunogène selon la revendication 10 caractérisée en ce qu'elle contient les peptides dont les formules correspondent à des séquences qui, dans les glycoprotéines d'enveloppe de HIV-2, SIV-1 et HIV-1 présentent une homologie en acides aminés supérieure à 50%.

15 33/ Composition immunogène selon l'une des revendications 31 ou 32, caractérisée en ce qu'elle contient au moins un peptide ou au moins un oligomère de ce peptide ou ce peptide sous forme conjuguée avec une molécule porteuse choisi parmi env4, env5, env6 et env10.

20 34/ Procédé de diagnostic in vitro de l'infection par HIV-2 dans un liquide biologique comprenant :

- la mise en contact de ce liquide biologique avec au moins un peptide selon l'une des revendications 1, 2, 3, 4, 5, 22 ou un conjugué de ces peptides avec une molécule porteuse ou des peptides gag ou env selon la revendication 26.

- la détection de la présence éventuelle d'un complexe antigène-anticorps par des méthodes physiques ou chimiques, dans ledit liquide biologique.

30 35/ Procédé de diagnostic in vitro de l'infection par HIV-2 dans un liquide biologique selon la revendication 34, caractérisé en ce que la détection du complexe antigène-anticorps éventuellement formé est réalisée grâce à des tests immunoenzymatiques (du type

51

ELISA) immunofluorescents (du type IFA) radioimmunologiques (du type RIA) ou des tests de radioimmunoprécipitation (du type RIPA).

36/ Kit pour le diagnostic in vitro de l'infection

5 par HIV-2 dans un liquide biologique caractérisé en ce qu'il comprend :

- une composition peptidique contenant un peptide selon l'une des revendications 1 à 5, 22, ou un mélange de ces peptides, ou un conjugué de ces peptides avec une molécule porteuse, ou les peptides gag ou env selon la

10 revendication 26,

- un réactif pour la constitution du milieu propice à la réalisation d'une réaction immunologique,

- un ou plusieurs réactifs éventuellement marqué pour la 15 détection du complexe antigène-anticorps formé par la réaction immunologique,

- un liquide biologique de référence dépourvu d'anticorps reconnus par la susdite composition peptidique.

20

25

30

35

FIG. 1.A

HIV2.ROD

R

```

GTCGCTCTGCGGAGAGGGCTGGCAGATTGAGCCCTGGGAGGTTCTCCAGCACTAGCAGG
TAGAGCCTGGGTGTTCCCTGCTAGACTCTCACCAAGCACTGGCCGGTGTGGCAGACGG
CCCCACCGCTTGCTTAAACCTCTTAATAAAAGCTGCCAGTTAGAAGCAAGTTAAGT
GTCTGCTCCCATCTCTCCTAGTCGCCCTGGTCATTGGTGTTCACCTGAGTAACAAGA
200
CCCTGGTCTGTTAGGACCCCTTGTGCTTGGAAACCGAGGCAGGAAAATCCCTAGCAGG
300
TTGGCCCTGAACAGGGACTTGAAAGAAAGACTGAGAAGTCTTGGAACACGGCTGAGTGAAG
GCAGTAACGGCGGCAGGAACAAACCACGACGGAGTGCTCCTAGAAAGGCCGGCCGAGG
400
CACCAAAGGCAGCGTGTGGAGCGGGAGGAGAACAGGCTCCGGGTGAAGGTAAGTACCTA
CACCAAAAATGTAGCCGAAAGGGCTTGCTATCCTAACAGGTAGAAGATTGTG
500
MetGlyAlaArgAsnSerValLeuArgGlyLysAlaAspGluLeuGluArgIle
GGAGATGGGCCCGAGAAACTCCGTCTTGAGAGGGAAAAAGCAGATGAATTAGAAAGAAT
600
ArgLeuArgProGlyGlyLysLysTyrArgLeuLysHisIleValTrpAlaAlaAsn
CAGGTTACGGCCCAGGAAAGAAAAACTACAGGCTAAACATATTGTGTGGCAGCGAA
LysLeuAspArgPheGlyLeuAlaGluSerLeuLeuGluSerLysGluGlyCysGlnLys
700
TAAATTGGACAGATTGGATTAGCAGAGAGCCCTGGCAGTCAAAAGAGGGTTGTCAAAA
IleLeuThrValLeuAspProMetValProThrGlySerGluAsnLeuLysSerLeuPhe
AATTCTTACAGTTTAGATCCAATGGTACCGACAGGTTCAAGAAAATTAAAAAGTUTTT
AsnThrValCysValIleTrpCysIleHisAlaGluGluLysValLysAspThrGluGly
TAATACTGTCTGGTCATTGGTGCATACACGCAGAAGAGAAAGTAAAGATACTGAAAG
800
AlaLysGlnIleValArgArgHisLeuValAlaGluThrGlyThrAlaGluLysMetPro
AGCAAAACAAATAGTGGGAGACATCTAGTGGCAGAAACAGGAACACTGCAGAGAAATGCC

```

FIG. 1A

2/35

SerThrSerArgProThrAlaProSerSerGluLysGlyGlyAsnTyrProValGlnHis
 AAGCACAAAGTAGACCAACAGCACCCTAGCGAGAAGGGAGGAAATTACCCAGTGCAACA
 ValGlyGlyAsnTyrThrHisIleProLeuSerProArgThrLeuAsnAlaTrpValLys
 TGTAGGCCAACTACACCCATATACCGCTGAGTCCCCGAACCTAAATGCCTGGGTAAA
 1000
 LeuValGluGluLysLysPheGlyAlaGluValValProGlyPheGlnAlaLeuSerGlu
 ATTAGTAGAGGAAAAAAAGTCCGGGCAGAAGTAGTGCAGGATTCAGGCAGTCTCAGA
 GlyCysThrProTyrAspIleAsnGlnMetLeuAsnCysValGlyAspHisGlnAlaAla
 AGGCTGCACGCCCTATGATATCAACCAAATGCTTAATTGTGTGGCGACCATCAAGCAGC
 1100
 MetGlnIleIleArgGluIleIleAsnGluGluAlaAlaGluTrpAspValGlnHisPro
 CATGCCAGATAATCAGGGAGATTATCAATGAGGAAGCAGCAGAATGGGATGTGCAACATCC
 1200
 IleProGlyProLeuProAlaGlyGlnLeuArgGluProArgGlySerAspIleAlaGly
 AATACCAGGCCCTTACCAAGGGGGAGCTTAGAGAGGCCAAGGGGACTGACATAGCAGG
 ThrThrSerThrValGluGluGlnIleGlnTrpMetPheArgProGlnAsnProValPro
 GACAACAAGCACAGTAGAACAGATCCAGTGGATGTTAGGCCACAAAATCCTGTACC
 1300
 ValGlyAsnIleTyrArgArgTrpIleGlnIleGlyLeuGlnLysCysValArgMetTyr
 AGTAGGAAACATCTATAGAACATGGATCCAGATAGGATTGCAGAAGTGTGTCAGGATGTA
 AsnProThrAsnIleLeuAspIleLysGlnGlyProLysGluProPheGlnSerTyrVal
 CAACCCGACCAACATCCTAGACATAAACAGGGACCAAGGGAGCCGTCCAAAGCTATGT
 1400
 AspArgPheTyrLysSerLeuArgAlaGluGlnThrAspProAlaValLysAsnTrpMet
 AGATAGATTCTACAAAGCTTGAGGGCAGAACAAACAGATCCAGCAGTGAAGAACATTGGAT
 1500
 ThrGlnThrLeuLeuValGlnAsnAlaAsnProAspCysLysLeuValLeuLysGlyLeu
 GACCCAAACACTGCTAGTACAAATGCCAACCCAGACTGTAAATTAGTGTCTAAAGGACT
 GlyMetAsnProThrLeuGluGluMetLeuThrAlaCysGlnGlyValGlyGlyProGly
 AGGGATGAACCCCTACCTTAGAACAGAGATGCTGACCGCTGTCAAGGGGTAGGTGGGCCAGG
 1600
 GlnLysAlaArgLeuMetAlaGluAlaLeuLysGluValIleGlyProAlaProIlePro
 CCAGAAAGCTAGATTAATGGCAGAGGCCCTGAAAGAGGTCTAGGACCTGCCCTATCCC
 PheAlaAlaAlaGlnGlnArgLysAlaPheLysCysTrpAsnCysGlyLysGluGlyHis
 ATTGGCAGGCCAGCAGCAGAGAACAGCATTAAATGCTGGAACTGTGGAAAGGGCA
 1700
 SerAlaArgGlnCysArgAlaProArgArgGlnGlyCysTrpLysCysGlyLysProGly
 CTGGCAAGACAATGCCGAGCACCTAGAACGGCAGGGCTGCTGGAAAGTGTGGTAAGCCAGG
 1800
 ThrGlyArgPhePheArgThrGlyProLeuGly
 HisIleMetThrAsnCysProAspArgGlnAlaGlyPheLeuGlyLeuGlyProTrpGly
 ACACATCATGACAAACTGCCAGATAGACAGGCAGGTTTTAGGACTGGGCCCTGGGG
 LysGluAlaProGlnLeuProArgGlyProSerSerAlaGlyAlaAspThrAsnSerThr
 LysLysProArgAsnPheProValAlaGlnValProGlnGlyLeuThrProThrAlaPro
 AAAGAAGCCCCGCAACTCCCCGTGGCCCAAGTCCGCAGGGCTGACACCAACAGCACC
 1900
 ProSerGlySerSerSerGlySerThrGlyGluIleTyrAlaAlaArgGluLysThrGlu
 ProValAspProAlaValAspLeuLeuGluLysTyrMetGlnGlnGlyLysArgGlnArg
 CCCAGTGGATCCAGCAGTGGATCTACTGGAGAAATATATGCAGCAAGGGAAAAGACAGAG
 ArgAlaGluArgGluThrIleGluGlySerAspArgGlyLeuThrAlaProArgAlaGly
 GluGlnArgGluArgProTyrLysGluValThrGluAspLeuLeuHisLeuGluGlnGly
 AGAGCAGAGAGAGAGACCATAAGGAAGTGACAGAGGACTTACTGCCACCTCGAGCAGGG
 (fig. 1A-suite 1)

3 / 35

GlyAspThrIleGlnGlyAlaThrAsnArgGlyLeuAlaAlaProGlnPheSerLeuTrp
GluThrProTyrArgGluProProThrGluAspLeuLeuLysLeuAsnSerLeuPheGly
GGAGACACCATAACAGGGAGCCACCAACAGAGGACTTGCTGCACCTCAATTCTCTCTTTGG
 2100
LysArgProValValThrAlaTyrIleGluGlyGlnProValGluValLeuLeuAspThr
LysAspGln
AAAAGACCAGTAGTCACAGCATACTGAGGGTCAGCCAGTAGAAGTCTGTAGACACA
 2200
GlyAlaAspAspSerIleValAlaGlyIleGluLeuGlyAsnAsnTyrSerProLysIle
GGGGCTGACGACTCAATAGTACCGAGGAATAGAGTTAGGGAAACAATTATAGCCCCAAAAATA
ValGlyGlyIleGlyGlyPheIleAsnThrLysGluTyrLysAsnValGluIleGluVal
CTAGGGGAATAGGGGATTCAAATACCAAGGAATATAAAATGTAGAAATAGAAGTT
LeuAsnLysLysValArgAlaThrIleMetThrGlyAspThrProIleAsnIlePheGly
CTAAATAAAAAGCTACGGGCCACCATAATGACAGGGCACCCCCAATCAACATTGGC
 2300
ArgAsnIleLeuThrAlaLeuGlyNetSerLeuAsnLeuProValAlaLysValGluPro
AGAAATATTCTGACAGCCTTAGGCATGTCATTAAATCTACCAGTCGCCAAAGTAGAGCCA
IleLysIleMetLeuLysProGlyLysAspGlyProLysLeuArgGlnTrpProLeuThr
ATAAAAATAATGCTAAAGCCAGGGAAAGATGGACCAAAACTGAGACAATGGCCCTAAC
LysGluLysIleGluAlaLeuLysGluIleCysGluLysMetGluLysGluGlyGlnLeu
AAAGAAAAAAATAGAACCACTAAAAGAAATCTGTGAAAAAATGGAAAAAGAAGGCCAGCTA
 2500
GluGluAlaProProThrAsnProTyrAsnThrProThrPheAlaIleLysLysLysAsp
GAGGAAGCACCTCCAACTAACTCCTATAATACCCCCACATTGCAATCAAGAAAAAGGAC
LysAsnLysTrpArgMetLeuIleAspPheArgGluLeuAsnLysValThrGlnAspPhe
AAAAACAAATGGAGGATGCTAATAGATTCAGAGAACTAAACAAGGTAACCTAACAGATTTC
 2600
ThrGluIleGlnLeuGlyIleProHisProAlaGlyLeuAlaLysLysArgArgIleThr
ACAGAAATTCTAGTTAGGAATTCCACACCCAGCAGGGTTGGCCAAGAAGAGAAGAATTACT
 2700
ValLeuAspValGlyAspAlaTyrPheSerIleProLeuHisGluAspPheArgProTyr
GTACTAGATGTAGGGATGCTTACTTTCCATACCACTACATGAGGACTTTAGACCATAT
 2800
ThrAlaPheThrLeuProSerValAsnAsnAlaGluProGlyLysArgTyrIleTyrLys
ACTGCCATTACTCTACCATCAGTGAACAATGCAGAACCAAGGAAAAAGATAACATATATAAA
ValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnHisThrMetArgGlnVal
GTCTGCCACAGGGATGGAAGGGATCACCAAGCAATTTCACACACACAATGAGACAGGTA
LeuGluProPheArgLysAlaAsnLysAspValIleIleIleGlnTyrMetAspAspIle
TTAGAACCAATTCTAGAAAAAGCAAACAAGGATGTCATTATCATTCACTGGATGATATC
 2900
LeuIleAlaSerAspArgThrAspLeuGluHisAspArgValValLeuGlnLeuLysGlu
CTAATAGCTAGTGACAGGACAGATTAGAACATGATAGGGTAGTCCTGCAGCTCAAGGAA
LeuLeuAsnGlyLeuGlyPheSerThrProAspGluLysPheGlnLysAspProProTyr
CTTCTAAATGGCCTAGGATTCTACCCCAGATGAGAAGTTCCAAAAAGACCCCTCCATAC
LysTrpMetGlyTyrGluLeuTrpProThrLysTrpLysLeuGlnLysIleGlnLeuPro
ACTGGATGGGCTATGAACTATGGCCAACATAATGGAAGTTGCAGAAAATACAGTTGCC
 3100
LysGluIleTrpThrValAsnAspIleGlnLysLeuValGlyValLeuAsnTrpAla
AAAAAGAAAATATGGACAGTCAATGACATCCAGAACGCTAGTGGGTGTCCTAAATTGGGCA

(fig.1A-suite 2)

4/35

AlaGlnLeuTyrProGlyIleLysThrLysHisLeuCysArgLeuIleArgGlyLysMet
 GCACAACTCTACCCAGGGATAAAACACCAAACACTTATGTAGGTTAACAGAGGAAAAATG
 3200
 ThrLeuThrGluGluValGlnTrpThrGluLeuAlaGluAlaGluLeuGluAsnArg
 ACACACTCACAGAAGAAGTACAGTGGACAGAATTAGCAGAAGCAGAGCTAGAAGAAAACAGA
 3300
 IleIleLeuSerGlnGluGlnGluGlyHisTyrTyrGlnGluGluLysGluLeuGluAla
 ATTATCCTAACGCCAGGAACAAGAGGGACACTATTACCAAGAAGAAAAGAGCTAGAACCA
 ThrValClnLysAspGluGluAsnGluTrpThrTyrLysIleHisGlnGluGluLysIle
 AGAGTCCAAAAGGATCAAGAGAATGAGTGGACATATAAAATACACCAGGAAGAAAAAATT
 LeuLysValGlyLysTyrAlaLysValLysAsnThrHisThrAsnGlyIleArgLeuLeu
 CTAAAAGTAGGAAAATATGCAAAGGTGAAAAACACCCATACCAATGGAATCAGATTGTTA
 AlaGlnValValGlnLysIleGlyLysGluAlaLeuValIleTrpGlyArgIleProLys
 GCACAGGTAGTTCAGAAAATAGGAAAAGAACGACTAGTCATTGGGGACGAATACCAAAA
 3500
 PheHisLeuProValGluArgGluIleTrpGluGlnTrpTrpAspAsnTyrTrpGlnVal
 TTTCACCTACCAGTAGAGAGAGAAATCTGGGAGCAGTGGTGGGATAACTACTGGCAAGTG
 3600
 ThrTrpIleProAspTrpAspPheValSerThrProProLeuValArgLeuAlaPheAsn
 ACATGGATCCCAGACTGGGACTTCGTGCTACCCCACCACTGGTCAGGTTAGCGTTAAC
 LeuValGlyAspProIleProGlyAlaGluThrPheTyrThrAspGlySerCysAsnArg
 CTGGTAGGGGATCCTATACCAAGGTGCAGAGACCTCTACACAGATGGATCCTGCAATAGG
 3700
 GlnSerLysGluGlyLysAlaGlyTyrValThrAspArgGlyLysAspLysValLysLys
 CAATCAAAAGAAGGAAAAGCAGGATATGTAACAGATAGAGGGAAAGACAAGTAAAGAAA
 LeuGluGlnThrThrAsnGlnAlaGluLeuGluAlaPheAlaMetAlaLeuThrAsp
 CTAGAGCAAACATCCAATCAGCAAGCAGAACTAGAACGCCTTGCATGGCACTAACAGAC
 3800
 SerGlyProLysValAsnIleIleValAspSerGlnTyrValMetGlyIleSerAlaSer
 TCGGGTCCAAAAGTTAATATTATAGACTCACAGTATGTAATGGGATCAGTCCAAGC
 3900
 GlnProThrGluSerGluSerLysIleValAsnGlnIleIleGluGluMetIleLysLys
 CAACCAACAGAGTCAGAAAGTAAATAGTGAACCAGATCATAGAAGAAATGATAAAAAAG
 GluAlaIleTyrValAlaTrpValProAlaHisLysGlyIleGlyGlyAsnGluVal
 GAAGCAATCTATGTTGCATGGGTCCCAGGCCACAAAGGCATAGGGGGAAACCAGGAAGTA
 4000
 AspHisLeuValSerGlnGlyIleArgGlnValLeuPheLeuGluLysIleGluProAla
 GATCATTAGTCAGGGTATCAGACAAGTGTGTTCCCTGGAAAAAAATAGAGCCCGCT
 GlnGluGluHisGluLysTyrHisSerAsnValLysGluLeuSerHisLysPheGlyIle
 CAGGAAGAACATGAAAATATCATAGCAATGTAAAAGAACTGTCTCATAAATTGGAATA
 4100
 ProAsnLeuValAlaArgGlnIleValAsnSerCysAlaGlnCysGlnGlnLysGlyGlu
 CCCAATTTAGTGGCAAGGCAAATAGTAAACTCATGTGCCAATGTCAACAGAAAGGGGAA
 4200
 AlaIleHisGlyGlnValAsnAlaGluLeuGlyThrTrpGlnMetAspCysThrHisLeu
 GCTATACATGGCAAGTAAATGCAGAACTAGGCACCTGGCAAATGGACTGGCACACATTAA
 GluGlyLysIleIleIleValAlaValAlaSerGlyPheIleGluAlaGluVal
 GAAGGAAAGATCATTATAGTAGCAGTACATGTTGCAAGTGGATTATAGAAGCAGAAGTC
 4300
 IleProGlnGluSerGlyArgGlnThrAlaLeuPheLeuLysLeuAlaSerArgTrp
 ATCCCACAGGAATCAGGAAGACAAACAGCAGTCTTCCATTGAAACTGGCAAGTAGGTGG
 (fig.1A-suite 3)

6/35

AlaGluValLeuGluIleLeuAla
 GlnArgSerTrpArgTyrTrpHisAspGluGlnGlyMetSerGluSerTyrThrLysTyr
 GGCAGAGGTCTGGAGATACTGGCATGATGAACAAGGGATGTCAGAAAGTTACACAAAGT
 5500
 ArgTyrLeuCysIleIleGlnLysAlaValTyrMetHisValArgLysGlyCysThrCys
 ATAGATATTGTGCATAATACAGAAAGCAGTGTACATGCATGTTAGGAAAGGGTGTACTT
 LeuGlyArgGlyHisGlyProGlyGlyTrpArgProGlyProProProProPro
 GCCTGGGGAGGGGACATGGCCAGGAGGGTGGAGACCAGGCCCTCCTCCTCCCCCTC
 5600
 MetAlaGluAlaProThrGluLeuProProValAspGlyThrProLeu
 GlyLeuVal***
 CAGGTCTGGCTAATGGCTGAAGCACCAACAGAGCTCCCCCGTGGATGGGACCCCAC
 ArgGluProGlyAspGluTrpIleIleGluIleLeuArgGluIleLysGluGluAlaLeu
 GAGGCAGCCAGGGATGAGTGGATAATAGAAATCTTGAGAGAAATAAAAGAACAGCTT
 LysHisPheAspProArgLeuLeuIleAlaLeuGlyLysTyrIleTyrThrArgHisGly
 MetGlu
 AAAGCATTGACCCCTCGCTGCTAATTGCTCTGGCAAATATATCTACTAGACATGG
 5800
 AspThrLeuGluGlyAlaArgGluLeuIleLysValLeuGlnArgAlaLeuPheThrHis
 ThrProLeuLysAlaProGluSerSerLeuLysSerCysAsnGluProPheSerArgThr
 AGACACCCTGAAGGGGCCAGAGAGCTCATTAAAGTCCCTGCAACGAGCCCTTCACGCA
 PheArgAlaGlyCysGlyHisSerArgIleGlyGlnThrArgGlyGlyAsnProLeuSer
 SerGluGlnAspValAlaThrGlnGluLeuAlaArgGlnGlyGluGluIleLeuSerGln
 CTTCAGAGCAGGATGTGGCCACTCAAGAATTGGCCAGACAAGGGAGGAAATCCTCTCTC
 5900
 AlaIleProThrProArgAsnMetGln
 LeuTyrArgProLeuGluThrCysAsnAsnSerCysTyrCysLysArgCysCysTyrHis
 AGCTATACCGACCCCTAGAAACATGCAATAACTCATGCTATTGTAAGCGATGCTGCTACC
 6000
 MetAsnGluArgAlaAsp
 CysGluMetCysPheLeuAsnLysGlyLeuGlyIleCysTyrGluArgLysGlyArgArg
 ATTGTCAGATGTGTTCTAAACAAGGGCTGGGATATGTTATGAACGAAAGGGCAGAC
 GluGluGlyLeuGlnArgLysLeuArgLeuIleArgLeuLeuHisGlnThrSerGluTyr
 Met
 ArgArgThrProLysLysThrLysThrHisProSerProThrProAspLys
 GAACAAGGACTCCAAAGAAAATAAGACTCATCCGTCTCCTACACCAGACAAGTGAGTAT
 6100
 AspGluSerAlaAlaTyrCysHisPheIleSer
 MetAsnGlnLeuLeuIleAlaIleLeuLeuAlaSerAlaCysLeuValTyrCysThrGln
 GATGAATCAGCTGCTTATTGCCATTAGCTAGTGCTGCTTAGTATATTGCACCCA
 TyrValThrValPheTyrGlyValProThrTrpLysAsnAlaThrIleProLeuPheCys
 ATATGTAAGTGTCTATGGCGTACCCACGTGGAAAAATGCAACCATTCCCTCTTTG
 6200
 AlaThrArgAsnArgAspThrTrpGlyThrIleGlnCysLeuProAspAsnAspAspTyr
 TGCAACCAGAAATAGGGATACTTGGGAACCATACTGCTTGCTGACAATGATGATTA
 6300
 GlnGluIleThrLeuAsnValThrGluAlaPheAspAlaTrpAsnAsnThrValThrGlu
 TCAGGAAATAACTTGAATGTAACAGAGGCTTGTGATGCCATGGAATAATACAGTAACAGA
 GlnAlaIleGluAspValTrpHisLeuPheGluThrSerIleLysProCysValLysLeu
 ACAAGCAATAGAAGATGTCTGGCATCTATTGAGACATCAATAAAACCATGTGTCAAAC
 6400
 *(fig.1A-suite 5)

7/35

ThrProLeuCysValAlaMetLysCysSerSerThrGluSerSerThrGlyAsnAsnThr
AACACCTTATGTGTAGCAATGAAATGCAGCACAGAGAGCACCACAGGGAAACAACAC

ThrSerLysSerThrSerThrThrThrProThrAspGlnGluGlnGluIleSer
AACCTCAAAGAGCACAAGCACAACCACAACCACACCCCACAGACCAGGAGCAAGAGATAAG
6500

GluAspThrProCysAlaArgAlaAspAsnCysSerGlyLeuGlyGluGluGluThrIle
TGAGGATACTCCATGCGCACGCGAGACAACGTGCTCAGGATTGGGAGAGGAAGAAACGAT
6600

AsnCysGlnPheAsnMetThrGlyLeuGluArgAspLysLysLysGlnTyrAsnGluThr
CAATTGCCAGTTCAATATGACAGGATTAGAAACAGATAAGAAAAACAGTATAATGAAAC

TrpTyrSerLysAspValValCysGluThrAsnAsnSerThrAsnGlnThrGlnCysTyr
ATGGTACTCAAAAGATGTGGTTGTGAGACAAATAATGCACAAATCAGACCCAGTGTAA
6700

MetAsnHisCysAsnThrSerValIleThrGluSerCysAspLysHisTyrTrpAspAla
CATGAACCATTGCAACACATCAGTCATCACAGAATCATGTGACAAGCACTATTGGGATGC

IleArgPheArgTyrCysAlaProProGlyTyrAlaLeuLeuArgCysAsnAspThrAsn
TATAAGGTTAGATACTGTGCACCACCGGGTATGCCCTATTAAAGATGTAATGATAACCAA
6800

TyrSerGlyPheAlaProAsnCysSerLysValValAlaSerThrCysThrArgMetMet
TTATTCAAGGCTTGCACCCAACTGTTCTAAAGTAGTAGCTTCTACATGCACCAGGATGAT
6900

GluThrGlnThrSerThrTrpPheGlyPheAsnGlyThrArgAlaGluAsnArgThrTyr
GGAAACGCAAACCTCCACATGGTTGGCTTAATGGCACTAGAGCAGAGAATAGAACATA

IleTyrTrpHisGlyArgAspAsnArgThrIleIleSerLeuAsnLysTyrTyrAsnLeu
TATCTATTGGCATGGCAGAGATAATAGAACTATCATCAGCTTAAACAAATATTATAATCT
7000

SerLeuHisCysLysArgProGlyAsnLysThrValLysGlnIleMetLeuMetSerGly
CAGTTGCATTGTAAGAGGCCAGGGATAAGACAGTGAACAAATAATGCTTATGTCAGG

HisValPheHisSerHisTyrGlnProIleAsnLysArgProArgGlnAlaTrpCysTrp
ACATGTGTTCACTCCCACTACCAGCCGATCAATAAGACCCAGACAAGCATGGTGCTG
7100

PheLysGlyLysTrpLysAspAlaMetGlnGluValLysGluThrLeuAlaLysHisPro
GTTCAAAGGCAAATGGAAAGACGCCATGCAGGAGGTGAAGGAAACCCCTGCAAAACATCC
7200

ArgTyrArgGlyThrAsnAspThrArgAsnIleSerPheAlaAlaProGlyLysGlySer
CAGGTATAGAGGAACCAATGACACAAGGAATATTAGCTTGCAGCGCCAGGAAAGGCTC

AspProGluValAlaTyrMetTrpThrAsnCysArgGlyGluPheLeuTyrCysAsnMet
AGACCCAGAAGTAGCATACTGTGGACTAACTGCAGAGGAGTTCTACTGCAACAT
7300

ThrTrpPheLeuAsnTrpIleGluAsnLysThrHisArgAsnTyrAlaProCysHisIle
GACTTGGTTCTCAATTGGATAGAGAATAAGACACACCGCAATTATGCCACCGTGCATAT

LysGlnIleIleAsnThrTrpHisLysValGlyArgAsnValTyrLeuProProArgGlu
AAAGCAAATAATTAAACACATGGCATAAGGTAGGGAGAAATGTATATTGCCCTCCCAGGGA
7400

GlyGluLeuSerCysAsnSerThrValThrSerIleIleAlaAsnIleAspTrpGlnAsn
AGGGGAGCTGCTGCAACTAACAGTAACCAGCATAATTGCTAACATTGACTGGCAAAA
7500

AsnAsnGlnThrAsnIleThrPheSerAlaGluValAlaGluLeuTyrArgLeuGluLeu
CAATAATCAGACAAACATTACCTTACTGGCAGAGGTGGCAGAACTATAAGATTGGAGTT

GlyAspTyrLysLeuValGluIleThrProIleGlyPheAlaProThrLysGluLysArg
GGGAGATTATAAATTGGTAGAAATAACCCAATTGGCTTGCACCTACAAAAGAAAAAG
7600

(fig.1A-suite 6)

8 / 35

9 / 35

GlutheLeuAlaGlyAlaCysArgGlyLeuTrpArgValLeuGluArgIleGlyArgGly
ArgLeuLeuArgAlaArgAlaGlyAlaCysGlyGlyTyrTrpAsnGluSerGlyGlyGlu
AGAGACTCTGGGGCGCGTGCAGGGCTTGTGGAGGGTATTGGAACGAATCCGGAGGGG
8600
IleLeuAlaValProArgArgIleArgGlnGlyAlaGluIleAlaLeuLeu
TyrSerArgPheGlnGluGlySerAspArgGluGlnLysSerProSerCysGluGlyArg
AATACTCGCGGTTCCAAGAACGGATCAGACAGGGAGCAGAAATGCCCTCCTGTGAGGGAC
8700
GlnTyrGlnGlnGlyAspPheMetAsnThrProTrpLysAspProAlaAlaGluArgGlu
GGCAGTATCAGCAGGGAGCTTATGAATACTCCATGGAAGGACCCAGCAGCAAAGGG
LysAsnLeuTyrArgGlnGlnAsnMetAspAspValAspSerAspAspAspAspGlnVal
AGAAAAATTGTACAGGCAACAAAATATGGATGATGTAGATTGATGATGACCAAG
8800
ArgValSerValThrProLysValProLeuArgProMetThrHisArgLeuAlaIleAsp
TAAGAGTTCTGTCACACCAAAAGTACCAACTAACGACCAATGACACATAGATTGGCAATAG
MetSerHisLeuIleLysThrArgGlyLeuGluGlyMetPheTyrSerGluArgArg
ATATGTCACATTTAATAAAAACAAGGGGGACTGGAAGGGATGTTACAGTGAAGAA
8900
HisLysIleLeuAsnIleTyrLeuGluLysGluGluGlyIleIleAlaAspTrpGlnAsn
GACATAAAATCTAAATATACTTAGAAAAGGAAGGAAGGGATAATTGCAGATTGGCAGA
9000
TyrThrHisGlyProGlyValArgTyrProMetPhePheGlyTrpLeuTrpLysLeuVal
ACTACACTCATGGGCCAGGAGTAAGATAACCAATGTTCTTGGGTGGCTATGGAAGCTAG
ProValAspValProGlnGluGlyGluAspThrGluThrHisCysLeuValHisProAla
TACCACTAGATGTCCCACAAGAACGGGAGGACACTGAGACTCACTGCTTAGTACATCCAG
9100
GlnThrSerLysPheAspAspProHisGlyGluThrLeuValTrpGluPheAspProLeu
CACAAACAAGCAAGTTGATGACCCGCATGGGAGACACTAGTCTGGAGTTGATCCCT
LeuAlaTyrSerTyrGluAlaPheIleArgTyrProGluGluPheGlyHisLysSerGly
TGCTGGCTTATAGTTACGAGGCTTTATTGGTACCCAGAGGAATTGGGCACAAGTCAG
9200
LeuProGluGluGluTrpLysAlaArgLeuLysAlaArgGlyIleProPheSer
GCCTGCCAGAGGAAGAGTGGAAAGGCGAGACTGAAAGCAAGAGGAATACCATTTAGTTAAA
9300
GACAGGAACAGCTATACTTGGTCAGGGCAGGAAGTAACAACTAACAGAAACAGCTGAGACTGC
AGGGACTTCCAGAACGGGCTGTAACCAAGGGAGGGACATGGGAGGAGCTGGTGGGAAC
9400
GCCCTCATATCTCTGTATAAAATACCCGCTAGCTTGCATTGTACTTCGGTCGCTCTGC
GGAGAGGGCTGGCAGATTGAGGCCCTGGGAGGTTCTCCAGCAGTAGCAGGTAGAGCCTGG
9500
GTGTTCCCTGCTAGACTCTACCCAGCACTTGGCCGGTGCTGGGAGACGGCCCCACGCTT
9600
GCTTGCTTAAAAACCTCCTTAATAAGCTGCCAGTTAGAAGCA

(fig.1A-suite 8)

10/35

FIG 1B

AGTCGCTCTGGAGAGGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAC
 GTAGAGCCTGGGTGTTCCCTGCTAGACTCTACCCAGCACCTGGCCGGTCTGGCAGACT
 100
 GGCTCCACGCCCTGCTTAAAGACCTCTCAATAAAAGCTGCCATTAGAAGTAAGCTA
 CTGTGTGTTCCCATCTCTCTAGTCGCCGCTGGTCAACTCGGTACTCGGTAAATAAAAAG
 200
 ACCCTGGCTGTTAGGACCCCTGGTCTGTTAGGACCCCTCTGCTTGGAAACCGAAGCA
 300
 GGAAAATCCCTAGCAGATTGGCCCGAACAGGGACTTGAAGGAGACTGAGAGACTCCTG
 AGTACGGCTGAGTGAAGGCAGTAAGGGCGGCAGGAACCAACCACGACGGAGTGCTCCTAG
 400
 AAAGGCCGGGTCGGTACCAAGACGGCGTGAGGAGCGGGGAGAGAAGAGGGCTCCGGTTG
 CAGGTAAGTGCAACACAAAAAGGAAATAGCTGTCTTTATCCAGGAAGGGATAATAAGAT
 500
 GAGDMETGLYAL AARGASNSERVALLEUSERGLYLYSLYSALAASPGLULEUGLU
 AGAGTGGAGATGGCGCGAGAAACTCCGTCTGTCAAGGAAGAAAGCAGATGAATTAGA
 600
 LYSILEARGLEUARGPROGLYGLYLYSLYSTYRMETLEULYSHISVALVALTRP
 AAAAATTAGACTACGACCCGGGAAAGAAAAAGTACATGTTGAAGCATGTAGTATGGC
 ALAASNGLULEUASPARGPHEGLYLEUALAGLUSERLEULEUGLUASNLYSGLUGLYCYS
 AGCAAATGAATTAGATAGATTGGATTAGCAGAAAGCCTGGAGAACAAGAAGGATG
 700
 GLNLYSILELEUSERVALLEUALAPROLEUVALPROTHRLYSERGLUASNLEULYSSER
 TCAAAAATACTTCGGTCTTAGCTCCATTAGTGCCAACAGGCTCAGAAAATTAAAAAG
 LEUTYRASNTHRVALCYSVALILETRPCYSILEHISALAGLUGLULYSALLYSHISTR
 CCTTTATAACTGTCTGGTCATCTGGTCATTACCGCAGAAAGAGAAAGTGAACACAC
 800
 GLUGLUALALYSGLNILEVALGLNARGHISLEUVALMEGLUTHRLYTHRALAGLUTHR
 TGAGGAAGCAAAACAGATAGTGCAGAGACACCTAGTGTGGAAACAGGAACAGCAGAAAC
 900
 METPROLYSTHRSERARGPROTHRALAPROPHESERGLYARGGLYGLYASNTYRPROVAL
 TATGCCAAAAACAAGTAGACCAACAGCACCATTAGCGGCAGAGGAGGAAATTACCCAGT
 GLNGLNILEGLYGLYASNTYRTHRHSILEUPDLEUSERPROARGTHRLEUASNALATRP
 ACAACAAAATAGGTGGTAACTATACCCACCTACCAAGCCCCGAGAACATTAAATGCC
 1000
 VALLYSLIEGLUGLULYSLYSPHEGLYALAGLUGLULALVALSERGLYPHEGLNALALEU
 GGTAAAATTAAATAGAGGAGAAGAAATTGGAGCAGAAGTAGTGTCAAGGATTCAAGGCACT
 SERGLUGLYCYSLEUPROTYRASPILAEASGLNMETLEUASNCYSVALGLYASPHISGLN
 GTCAGAAGGCTGCCTCCCTATGACATTAATCAGATGTTAAATTGTGTGGAGACCATCA
 1100
 ALAALAMETGLNILEILEARGASPILEILEASGLUGLULALALAASPTRPASPLEUGLN
 AGCGGCTATGCAGATCATCAGAGATATTATAATGAGGAGGCTGCAGATTGGACTTGCA
 1200
 HISPROGLNGLNALAPROGLNGLYGLNLEUARGGLUPROSERGLYSERASPILEALA
 GCACCCACAACAAGCTCCACAACAAGGACAGCTAGGGAGCCGTCAAGGATCAGATATTGC
 GLYTHRTHRSERTHRVALGLUGLUGLNILEGLNTRPMETTYRARGGLNGLNASNP
 AGGAACAACAGTACAGTAGAAGAACAAATCCAGTGGATGTACAGACAACAGAACCCAT
 1300

FIG. 1B-

11/35

PROVALGLYASNILETYRARGARGTRPILEGLNLEUGLYLEUGLNLYSCYSVALARGMET
 ACCAGTAGGCAACATTACAGGAGATGGATCCAACCTGGGTGCCAAAATGTGTCAGAAAT
 TYRASNPROTHRASNILEUASPVALYSGLNGLYPROLYSGLUPROPHEGLNSERTYR
 GTATAACCCAACAAACATTCTAGATGAAAACAAGGCCAAAAGAGCCATTCAAGCTA
 1400
 VALASPARGPHETYRLYSSERLEUARGALAGLUGLNTHRASPPROALAYALLYSASNTRP
 TGTAGACAGGTTCTACAAAAGTTAACAGCAGAACAAACAGATCCAGCAGTAAAGAATTG
 1500
 METTHRGLNTHRLEULEUILEGLNASNALAASNPROASPCYSLYSLEUVALLEYSLY
 GATGACTCAAACACTGCTGATTCAAATGCTAACCCAGATTGCAAGCTAGTGCTGAAGGG
 LEUGLYTHRASNPROTHRLEUGLUGLUMETLEUTHRALACYSGLNGLYVALGLYGLYPRO
 GCTGGGTACGAATCCCACCCCTAGAAGAAATGCTGACGGCTGTCAAGGAGTAGGGGGCC
 1600
 GLYGLNLYSALAARGLEUMETALAGLUALALEULYSGLUALALEUALAPROALAPROILE
 AGGACAGAAGGCTAGATTAATGCCAGAACAGCCCTGAAAGAGGCCCTCCGACCAGGCCAAT
 POLVALLEUGLULEUTRP
 PROPHEALAAALAGLNGLNLYSGLYPROARGLYSPROILELYSCYSTRPASNCYSGLY
 CCCTTTGCAGCAGCCAACAGAAGGGACCAAGAAAGCCAATTAAAGTGTGGAATTGTGG
 1700
 GLUGLYARGTHRLEUCYSLYSALAMETGLNSERPROLYSLYSTHRGLYMETLEUGLUMET
 LYSGLUGLYHISSERALAARGGLNCYSARGALAPROARGARGGLNGLYCYSTRPLYSYS
 GAAGGAAGGACACTCTGCAAGGCAATGCAGAGCCCCAAGAACAGACAGOGATGCTGGAAATG
 1800
 TRPLYSASNGLYPROCYSTYRGLYGLNMETPROLYSGLNTHRGLYGLYPHEPHEARGPRO
 GLYLYSMETASPHISVALMETALALYSCYSPROASNARGGLNALAGLYPHELEUGLYLEU
 TGGAAAAATGGACCATGTTATGCCAAATGCCAAACAGACAGCAGGGCTTTAGGCCT
 TRPPROLEUGLYLYSGLUALAPROGLNPHEPROHISGLYSERSERALASERGLYALAASP
 GLYPROTRPGLYLYSLYSPROARGASNPHEPROMETALAGLNVALTHISGLNGLYLEUTHR
 TGGCCCTTGGGAAAGAACAGCCCCCAATTCCCCATGGCTCAAGTCATCAGGGCTGAC
 1900
 ALAASNCYSERPROARGARGTHRSERCYSGLYSERALALYSGLULEUHISALALEUGLY
 PROTHRALARPROPROGLUGLUPROALAVALASPLEULEULYSAASNTRMETHISLEUGLY
 GCCAACTGCTCCCCAGAACAGCTGTGGATCTGCTAAAGAACTACATGCACTTGG
 GLNALAALAGLUARGLYSGLNARGGLUALALEUGLNGLYGLYASPARGGLYPHEALAALA
 LYSGLNGLNARGGLUSERARGGLYLYSPROTYRLYSGLUVALTHRGLUASPLEULEUHIS
 CAACCGAGAGAGAACAGAGGGAAAGCCTTACAAGGAGGTGACAGAGGATTGCTGCA
 2000
 PROGLNPHESERLEUTRPARGARGPROVALYALTHRALARHISILEGLUGLYGLNPROVAL
 LEUASNRSERLEUPHEGLYGLYASPGLN
 CCTCAATTCTCTTTGGAGGAGACCAGTAGTCAGTCACTGCTCATATTGAAGGACAGCCTGTA
 2100
 GLUVALLEULEUASPTHRGLYALAASPASP SERILEVALTHRGLYILEGLULEUGLYPRO
 GAAGTATTATTAGATACAGGGGCTGATGATTCTATTGTAACAGGAATAGAGTTAGGTCCA
 HISTYRTHRPROLYSILEYALGLYGLYILEGLYGLYPHEILEASNTHRRLYSGLUTYRLYS
 CATTATACCCAAAAATAGTAGGAGGAATAGGAGGTTTTATTAAACTAAAGAACACAAA
 2200
 ASNVALGLUILEGLUVALLEUGLYLYSARGILELYSGLYTHRILEMETTHRGLYASPTHR
 AATGTAGAAATAGAAGTTTAGGCAAAAGGATTAAAGGGACAATCATGACAGGGACACC
 PROILEASNILEPHEGLYARGASNLEULEUHLRALALEUGLYNETSERLEUASNLEUPRO
 CCCGATTAACATTTGGTAGAAATTTACTAACAGCTCTGGGATGTCCTAAATCTTCCC
 2300
 ILEALALYSVALGLUPROVALLYSERPROLEULYSPROGLYLYSASPGLYPROLYSLEU
 ATAGCTAAGGTAGAGCCTGAAAGTCGCCCTAAAGCCAGGAAAGGATGGACCAAAATTG
 2400
 LYSGLNTRP PROLEUSERLYSGLULYSSILEYALAEALEUARGGLUJLECYSGLULYSMET
 AACAGTGGCCATTATCAAAAGAAAAGATAGTTGCATTAAGAGAAATCTGTGAAAAGATG

(fig.1B-suite 1)

12/35

GLULYSASPGLYGLNLEUGLUGLUALAPROPROTHRASNPROTYRASNTHRPROTHRPH
GAAAAAGATGGTCAGTTGGAGGAAGCTCCCCCACCAATCCATATAAACACCCCCACATT
2500
ALAILELYSLYSLYSASPLYSASNLYSSTRPARGMETLEUILEASPPHEARGGLULEUASN
GCTATAAAGAAAAAGGATAAAAACAAATGGAGAATGCTGATAGATTAGGGAACTAAAT
ARGVALTHRGLNASPPHETHRGLUVALGLNLEUGLYILEPROHISPROALAGLYLEUALA
AGGGTCACTCAAGACTTTACCGAAGTCCAATTAGGAATACCACACCCCTCCAGGACTAGCA
2600
LYSARGLYSARGILETHRVALLEUASPILEGLYASPALATYRPHESERILEPROLEUASP
AAAAGGAAAAGGATTACAGTACTGGATATAGGTGACGCATATTCTCTACACCTCTAGAT
2700
GLUGLUPHEARGGLNTYRTHRALAPHETHRLEUPROSERVALASNALSAGLUPROGLY
GAAGAATTAGGCAGTACACTGCCTTACTTACCATCAGTAAATAATGCCAGAGCCAGGA
LYSARGTYRILETYRLYSVALLEUUPROGLNGLYTRPLYSGLYSERPROALAILEPHEGLN
AAACGATACATTATAAGGTTCTGCCTCAGGATGGAAGGGTCACCAGCCATCTCCAA
2800
TYRTHRMETARGHISVALLEUUGLUPROPHEARGLYSALAASNPROASPVALTHRLEUVAL
TACACTATGAGACATGTGCTAGAACCTTCAGGAAGGCAAATCCAGATGTGACCTTAGTC
GLNTYRMETASPASPILEUILEALASERAASPARGTHRASPLEUGLUHISASPARGV
CACTATATGGATGACATCTTAATAGCTAGTGACAGGACAGACCTGGAACATGACAGGGTA
2900
VALLEUGLNLEULYSGLULEULEUASNRSERILEGLYPHESESERPROGLUGLULYSPE
GTTTACAGTTAAAAGAACTCTTAAATAGCATAGGGTTCTCCAGGACAAATGGAAAGTTG
3000
GLNLYSASPPROPROPHEGLNTRPMETGLYTYRGLULEUTRPPROTHRLYSSTRPLYSLEU
CAAAAGATCCCCCATTCAATGGATGGGTACGAATTGTGGCCGACAAATGGAAAGTTG
GLNLYSILEGLULEUPROGLNARGGLUTHRTRPVALASNASPILEGNLYSLEUVAL
CAAAAGATAGAGTTGCCACAAAGAGAGACCTGGACAGTGAATGATATACAGAAGTTAGTA
3100
GLYVALLEUASNTRPALAALAGLNILETYRPROGLYILELYSTHRLYSHISLEUCYSARG
GGAGTATTAAATTGGGCAGCTCAAATTATCCAGGTATAAAACAAACATCTCTGTAGG
LEUILEARGLYLYSMETTHRLEUTHRGLUGLVALGLNTRPTHRGLUMETALAGLUALA
TTAATTAGGGAAAAATGACTCTAACAGAGGAAGTTCACTGGACTGAGATGGCAGAAGCA
3200
GLUTYRGLUGLUASNLYSILEILELEUSERGLNGLUGLNGLUGLYCYSTYRTYRGLNGLU
GAATATGAGGAAAATAAAATAATTCTCAGTCAGGAACAAGAAGGATGTTTACCAAGAA
3300
SERLYSPROLEUGLUALATHRVALILELYSSERGLNASPASNGLNTRPSERTYRLYSILE
AGCAAGCCATTAGAACGCCACGGTGATAAAAGAGTCAGGACAATCACTGGTCTTATAAAATT
HISGLNGLUASPLYSILEEULYSVALGLYLYSPHEALALYSILELYSASNTHRHI
CACCAAGAACAAAATACTGAAAGTAGGAAAATTGCAAAGATAAAAGAACATACACATACC
3400
ASNGLYVALARGLEULEUALAHISVALILEGLNLYSILEGLYLYSGLUALAILEVALILE
AATGGAGTTAGACTATTAGCACATGTAATACAGAAAATAGGAAAGGAAGCAATAGTGATC
TRPGLYGLNVALPROLYSPHEHISLEUPROYALGLULYSASPVALTRPGLUGLNRPTRP
TGGGGACAGGTCCAAAATTCCACTTACCTACAGTTGAGAAGGATGTATGGAACAGTCGTGG
3500
THRASPTYRTRPGLNVALTHRTRPILEPROGLUTRPASPPHEILESERTHRPROPROLEU
ACAGACTATTGGCAGGTAACTGGATACCGGAATGGGATTTCATCTCAACACACCACATT
3600
VALARGLEUVALPHEASNLEUVALLYSASPPROILEGLUGLYGLUGLUTHRTYRTYRVAL
GTAAGATTAGTCCTCAATCTAGTGAAGGACCCCTATAGAGGGAGAAGAACCTATTATGTA
ASPGLYSERCYSSERLYSGLNSERLYSGLYLYSALAGLYTYRILETHRSPARGGLY
GATGGATCATGTAGTAAACAGTCAAAAGAACGAGGATATCACAGACAGACAGGGC

(fig. 1B - suite 2)

13/35

3700

LYSASPLYSVALLYSVALLEUGLUGLNTHRTHRASHNLNGLNALACLULEUGLUALAPHE
 AAAGACAAGGTAAAAGTGTAGAACAGACTACTAATCAACAAGCAGAATTGGAAGCATT
 LEUMETALALEUTHRASP SERGLYPROLYSALAASNILEILEVALASP SERCLNTYRVAL
 CTCATGGCATTGACAGACTCAGGCCAAAGGCAAATTATAGTAGACTACAATATGTT
 3800
 METGLYILEILETHRLYCYS PROTHRGLUSERGLUSERARGLEUVALASNGLNILEILE
 ATGGGAATAATAACAGGATGCCCTACAGAATCAGAGAGCAGGCTAGTTAACCAAATAATA
 3900
 GLUGLUMETILELYSLYTHRLGLUILETYRVALALAATRPVALPROALAHISLYSGLYILE
 GAAGAAATGATCAAAAAGACAGAAATTATGTGGCATGGTACCCAGCACACAAAGGTATA
 GLYGLYASNGLNGLUILEASPHISLEUVALSERGLNGLYILEARGGLNVALLEUPHELEU
 GGAGGAAACCAAGAAATAGACCACCTAGTTAGTCAGGGATTAGACAAGTTCTCTTCTG
 4000
 GLULYSILEGLUPROALAGNGLUGLUHISERLYSTYRHISERASHNILELYSGLULEU
 GAAAAGATAGAGCCAGCACAAAGAACATAGTAAATACCATAGTAAACATAAAAGAATTG
 VALPHELYSPEGLYLEUPROARGLEUVALALALYSGLNILEVALASP THRCYSASPLYS
 GTATTCAAATTGGATTACCCAGACTAGTGGCAAACAGATAGTACACATGTGATAAA
 4100
 CYSHISGLNLYSGLYGLUALAILEHISGLYGLNVALASN SERASPLEUGLYTHRTRPGLN
 TGT CATCAAAAAGGAGAAGCTACATGGCAGGTAAATT CAGACCTAAGGACTTGGCAA
 4200
 METASPCYSTHRHISLEUGLUGLYLISILEVALILEVALALAYLHISVALALASERGLY
 ATGGATTGTACCCATCTAGAGGGAAAAATAGTCATAGTTGCAGTACATGTAGCTAGTGGA
 PHEILEGLUALAGLUVALILEPROGLNGLUTHRLYARGGLNTHR ALALEUPHELEU
 TTCATAGAACAGTAATTCCACAAGAAACAGGAAGACAGACAGCACTATTCTGTAA
 4300
 LYSLEUALASERARGTRPPROILETHRHSLEUHIS THRASPASNGLYALAASN PHEALA
 AAATTGGCAAGCAGATGCCATTACACATCTGCACACAGATAATGGTGTAACTTTGCT
 SERGLNGLUVALLYSMETVALALATRPTRPALAGLYILEGLUHIS THR PHEGLYVALPRO
 TCGCAAGAAGTAAAGATGGTGCATGGTGGCAGGGATAGAGCACACCTTGGGTACCA
 4400
 TYRASNPROGLNSERGLNGLYVALVALGLUALAME TASHHISI SLEULYSASNGLNILE
 TACAATCCACAGAGTCAGGGAGTAGTGGAAAGCAATGAATCACCACCTGAAAATCAAATA
 4500
 ASPARGILEARGGLUGLNALAA NSERVALGLUTHRILEVALLEUMETALAYLHIS CYS
 GATAGAACAGCAAAATTCACTAGAAACCATAGTATTAAATGCCAGTTCTG
 METASNPHELYSARGRGGLYGLYILEGLYASPMETTHR PROALAGLUARGLEUILEASN
 ATGAATT TAAAGAAGGGAGGAATAGGGATATGACTCCAGCAGAAAGATTAAAC
 4600
 METILETHRTHRLUGLNGLUILELEG NLPHEGLNLSERLYSASN SERLYSPHELYSASH
 ATGATCACTACAGAACAAAGAAATACAATTCAACAATCAA AAAACTCAAATTTAAAAT
 PHEARGVALTYRTYRARGGLUGLYARGASPLNLEUTRPLYSGLYPROGLYGLULEU
 TTTCGGGTCTATTACAGAGAACGGCAGAGATCAGCTGTGGAGGGACCCGGTAGCTATTG
 4700
 TRPLYSGLYGLUGLYALAVALEU LYVALGLYTHRASPILEYSYVALVALPROARG
 TGGAAAGGGAGGAGGAGTCATCTAAAGGTAGGAACAGACATTAAGGTAGTACCCAGG
 4800
 ARGLYSALALYSILEILELYSASP TYRGLYGLYLYSGLUMETASP SERSER SERHIS
 QMETGLUGLUGLULY SARGTRP ILEVALVALPROTHR
 AGAAAGGCTAAATTATCAAAGATTATGGAGGAGGAAAAGAGATGGATAGTAGTTCCCAC
 METGLUASPLTHRGLYGLUALAARGGLUVALALA
 TRPARGILEPROGLUARGLEUGLUARGTRPHISERLEU ILELYSTYRLEU LYSTYR LYS
 ATGGAGGATACGGAGAGGCTAGAGAGGTGGCATGCCCTACAAAATATTGAAATATAA
 4900
 (fig. 1B - page 3)

THRLYSASPLEUCLNLYSALACYSTYRALPROHISIISLYSVALGLYTRPALATRPTPR
 AACTAAAGATCTACAAAAGGCTGCCTATGTGCCCATCATAAGGTGGATGGCATGGT
 THRCYSSERARGVALILEPHEPROLEUGLNGLUGLYSERHISLEUGLUVALGLNGLYTYR
 GACCTGCAGCAGAGTAATCTTCCCACTACAGGAAGGCCATTAGAAGTACAAGGGTA
 5000
 TRPASNLEUTHRPROGLUARGGLYTRPLEUSERTHRTHRTYRALAVALARGILETHRPTYR
 TTGGAATTGACACCAGAAAGACGGTGGCTCAGTACTTATGCAGTGAGGATAACCTGGT
 5100
 SERLYSASPHEPTRPTHRASPVALTHRPROGLUTYRALAASPILEULEUHISSETHR
 CTCAAAGGACTTTGGACAGATGTAACACCAGAATATGCAGATATTTACTGCATAGCAC
 TYRPHEPROCYSPEHTRALAGLYGLUVALARGARGLAILEARGGLYGLUARGLEULEU
 TTATTCCTGCTTACAGCGGGAGAAGTGAGAAGGGCCATCAGGGAGAACGACTGCT
 5200
 SERCYSCYSARGPHEPRODARGALAHISLYSHISGLNYALPROSERLEUGLNTYRLEUALA
 GTCTGCAGGTTCCAAGAGCTCATAAGCACCAAGGTACCAAGTCTACAGTACTTAGC
 LEUARGVALVALSERHISVALARGSERGLNGLYGLUASNPROTHRTRPLYSGLNTRPARG
 X METSERASPPROARGGLUARGILEPROPROGLYASNSERGLYGLU
 ACTGAGAGTAGTAAGTCATGTCAGATCCCAGGGAGAGAATCCCACCTGAAACACTGGAG
 5300
 ARGASPASNARGSERLEUARGVALALALYSGLNASNSERARGGLYASPLYSGLNARG
 GLUTHRILEGLYGLUALAPHEGLUTRPLEUASNARGTHRVALGLUGLUTLEASNARGGLU
 AAGAGACAATAGGAGAACGCCCTCGAGTGGCTAACAGAACAGTAGAGGGAGATAAACAGAG
 5400
 GLYGLYLYSPROPROTHRGLUGLYALAASNPHPROGLYLEUVALALYSVALLEUGLYILE
 ALAVALASNHSILEUPROARGGLULEUILEPHEGLNYALTRPGLNARGSERTRPGLUTYR
 AGGCCGTTAACCCACCTACCGAGGGAGCTAATTTCCAGGTTGGCAAAGGTCTGGAAAT
 LEUALA
 TRPHISASPGGLUGLNGLYMETSERGLNSERTYRTHRLYSTYRARGTYRLEUCYSLEUILE
 ACTGGCATGATGAACAAGGCATGTCACAAAGCTATACAAAATACAGATACTTGTGTTAA
 5500
 GLNLYSALALEUPHEMETHISCVSLSLYSGLYCYSARGCYSLEUGLYGLUGLYHISGLY
 TACAAAAGGCTTATTATGCATTGCAAGAAAGGCTGTAGATGTCTAGGGGAAGGACACG
 ALAGLYGLYTRPARGPROGLYPROPROPROPROPROPROGLYLEUALA R METGLU
 GGGCAGGGGGATGGAGACCAGGACCTCCTCCTCCCCCTCAGGACTAGCATAATGG
 5600
 GLUARGPROPROGLUASNGLUGLYPROGLNARGGLUPROTTPASPGGLUTRPTVALVALGLU
 AAGAAAGACCTCCAGAAAATGAAGGCCACAAAGGAAACCATGGATGAGTGGTAGTGG
 5700
 VALLEULYSGLULEULYSGLUGLUALALEULYSHISPHEASPPDARGLEULEUTHRALA
 AAGTTCTGAAAGAACTGAAAGAAGCTTAAAGCATTGATCCTCGGCTCTAACCG
 TATI METGLUTHRPROLEUARGGLUGLNGLUASNTER
 LEUGLYASNHSILETYRASNARGHISGLYASPTHRLEUGLUGLYALAGLYGLULEUILE
 CACTGGTAATCATATCTATAATAGACATGGAGACACCCTGAGGGAGCAGGAGAACTCA
 5800
 LEUGLUSERSERASNGLUARGSERSERTYRILESERGLUALAALAALAILEPROGLU
 ARGILELEUGLNARGALALEUPHEILEHISPHEARGSERGLYCYSSERHISSEARGILE
 TTAGAACCTCCAACGAGGGCTTCTACATTCAGAACGGGCTGCAGGCATTCCAGAA
 SERALAASNLEUGLYGLUGLUTLEUSERGLNLEUTYRARGPROLEUGLUALACYSTYR
 GLYGLNPROGLYGLYGLYASNPROLEUSERTHRILEPROPROSERARGSERMETLEU
 TCGGCCAACCTGGGGAGGAAATCCTCTCTCAACTATACCGCCCTCTAGAACGATGCTAT
 5900
 ASNTHRCYSTYRCVSLYSCVSCYSTYRHISCVSGLNPHECYPHELEULYSLYSGLY
 AACACATGCTATTGCAAAAGTGTGCTACCATTGCCAGTTTGTCTAAAAAGGGC
 6000
 LEUGLYILESERTYRGLULYSSERHISARGARGARGTHRPROLYSLYSALALYSALA
 ARTIMETARGSERHISTHRLYGLUGLUGLULEUARGARGCARGLEUARGLEU

(fig.1B-suite 4)

TTGGGATAAGTTATCACAACTCACACAGGAGAAGAACCTCCGAAGAAGGCTAAGGCT
 ASNTHR SERSERALASERASHGLU
 ILEHISLEULEUHISCLNTHRSERLYSTYRGLYLEUSERTRPLYSSERALAALATYRARG
 ENV METGLYCYSLEUGLYASHGLNLUEUILEALA
 AATACATCTTCTGCATCAAACGAGTAAGTATGGTTGTCTGAAATCAGCTGCTTATCG
 6100
 HISLEULEU
 ILECYSSERLYSCYSLEUTRPILEILECYSILECLNTYRALVALRVALPHETYRGLYVAL
 CCATCTGCTCTAAGTGTCTATGGATTATTGTATTCAATATGTCACAGTCTTTATGGT
 PROALATRPARGASHALATHRILEPROLEUPHECYSALATHRLYSASNARGASPTHRTRP
 TACCAGCTTGAGGAATGCGACAATTCCCCTCTGTGCAACCAAGAATAGGGACTT
 6200
 GLYTHRTHRGLNCYSLEUPROASPASNASPASPTYRSERGLULEUALAEUASNVALTHR
 GGGAAACAACCTCACTGCCCTACCAAGATAATGATGATTATTAGAATTGGCCCTTAATGTTA
 6300
 GLUSERPHEASPALATRPGLUASNTHRVALTHRGLUGLNLALATLEGLUASPVALTRPGLN
 CAGAAAGCTTGATGCTTGGGAGAATACAGTCACAGAACAGGAATAGAGGACGTATGGC
 LEUPHEGLUTHRSERILELYSPROCYSVALLYSLEUSERPROLEUCYSILETHRMEARG
 AACTCTTGAGACCTCAATAAGCCTTGTGAAAATTATCCCATTATGCATTACTATGA
 6400
 CYSAASNLYSSERGLUTHRASPLYSRPGLYLEUTHRLYSSERSERTHRTHRTHRALASER
 GATGCAATAAAAGTGAGACAGATAATGGGATTGACA AAAATCATCAACAACAGCAT
 THRTHRTHRTHRTHRTHRALALYSSERVALGLUTHRARGASPILEVALASNGLUTHRSER
 CAACAACAACAACACAGAAAATCACTAGAGACAAAGAGACATAGTCATGAGACTA
 6500
 PROCYSVALVALHISASPASNCYSTHRGLYLEUGLUGLNGLUPROMETILESERCYSLYS
 GTCTTGCTAGTTCTGATAATTGCACAGGTTGGAAACAGAGCCAAATGATAACCTGTA
 6600
 PHEASNMETTHRGLYLEEULYSARGASPLYSLYSGLUITYRASNGLUTHRTRPTYRSER
 AATTCAACATGACAGGGTTAAAAGAGACAAAGAAAAGGAGTACAATGAAACTTGGTACT
 ALAASPLEUVALCYSGLUGLNGLYASN SERTHRGLYASHNGLUSERARGCYSTRMETASN
 CTGCAGATCTGTTGTGAAACAAGGAATAGCACTGTAATGAAAGTAGATGTTACATGA
 6700
 HIS CYSASNTHR SERVAL ILEGlnGLUCYSCYSASPLYSASPTYRTRPASPALAILEARG
 ATCACTGTAATACTTCTGTTATCCAAGAGTGTGACAAAGATTATTGGGATGCTATTA
 CYSARGTYRCYSALAPROPROGLYTIRALALEULEUARGCYSASN ASPTHRASNTYRSER
 GATGTAGATATTGTGCACCTCCAGGTTATGCTTGTCTAGATGTAATGACACAAATTATT
 6800
 GLYPHEMETPROASNCYSERLYSVALVALVALSERSERCYSTHRARGMETMETGLUTHR
 CAGGCTTATGCCTAATGTTCTAAGGTAGTGGTCTCTCATGCACAAGGATGATGGAGA
 6900
 GLNTHRSERTHRTRPPHEARGPHEASNGLYTHRARGALAGLUASNARGTHR TYRILETYR
 CACAGACTCTACTTGGTTCCGGTTAATGGAACTAGAGCAGAAAATAGAACCTATATT
 TRPHISGLYARGASPAHNARGTHRILEILESERLEUASNLYSHISTYRASNLEUTHRME
 ACTGGCATGGTAGAGATAATAGGACTATAATTAGTCTAAATAAGCATTATAATCTAACAA
 7000
 LYS CYSARGARGPROGLYASHNLYSTHRALLEUROVALTHRILEMETSERALALEUVAL
 TGAAATGTAAGACCAAGGAAATAAGACAGTTTACCACTCACCATTATGTCTGCATTGG
 PHEHISERGLNPROVALASNGLUARGPROLYSGLNHALATRPCYSARGPHEGLYGLYASN
 TTTTCCACTCACAAACCAAGTCATGAGAGGGCAAAGCAGGCATGGTAGGTTGGAGGAA
 7100
 TRPLYSGLUALILELYSGLUVALLYSGLNTHRILEVALLYSHISPROARGTYRTHRGLY
 ATTGGAAAGGAGGCAATAAAAGAGGTGAAGCAGACCATGTCAAACATCCCAGGTATACTG
 7200
 THRASNASHRASPLYSILEASNLEUTHRALAPROARGGLYGLYASPPROGLUVALTHR
 GAACTAACAATACTGATAAAATCAATTGACGGCTCTAGAGGAGGAGATCCGGAAGTTA

(fig.1B-suite 5)

16/35

PHEMETTRPTHRASNCSARGGLYGLUPHELEUTYRCYSLYSMETASNTRPHELEUASN
 CCTTCATGTGGACAAATTGGAGAGGAGACTTCTCTACTGTAATGAATTGGTTCTAA
 7300
 TRPYALGLUASPARGSERLEUTHRTHRGLNLYS PROLYSGLUARCHISLYSARGASNTYR
 ATTCGGTAGAACATAGGAGCTAACTACCCAGAAGCCAAAGGAACGGCATAAAAGGAATT
 VALPROCYSHISILEARGGLNILEILEASNTHRTRPHISLYSVALGLYLYSASNVALTYR
 ACCTACCATGTCATATTAGAACAAATAATCAACACTTGGCATAAAAGTAGGCAAAATGTT
 7400
 LEUPROPROARGGLUGLYASPLEUTHRCYSASN SERTHRVALTHR SERLEU ILEALA ASN
 ATTTGCCTCCAAGAGAGGGAGACCTCACGTGTAACTCCACAGTGACCGACTCATAGCAA
 7500
 ILEASNTRPTHRASPGLYASNGLNTHR SERILETHR METSERALAGLUVALALAGLULEU
 ACATAAAATTGGACTGATGAAACCAAAACTAGTATCACCATGAGTGCAGAGGTGGCAGAAC
 TYRARGLEUGLULEUGLYASPTYRLYSLEUVALGLUILETHR PROILEGLYLEUALAPRO
 TGTATCGATTGGAATTGGAGATTATAAATTAGTAGAAATCACTCCAATTGGCTGGCCC
 7600
 THRASNVALYSARGTYRTHRTHRGLYGLYTHR SERARGASNLYSARGLYVALPHEVAL
 CCACAAATGTGAAGAGGTACACTACTGGTGGCACCTCAAGAAATAAAAGAGGGGTCTTG
 LEUGLYPHELEUGLYPHELEUALATHRALAGLYSERALAMETGLYALAA LASERLEUTHR
 TGCTAGGGTTCTGGGTTCTCGAACGGCAGGTTCTGCAATGGCGCGGCGTGGTGA
 7700
 VALTHR ALAGLN SERARGTHR LEU EULAGLYILEVALGLNGLNGLNGLNLEULEU
 CGGTGACCGCTCAGTCCCCACTTATTGGCTGGGATAGTGCAAGAACAGCAACAGCTGT
 7800
 ASPVALVALLYSARGGLNGLNGLULEUARGLEUTHRVALTRPGLYTHR LYS ASNLEU
 TGGACGTGGTCAAGAGACAACAAGAATTGTTGGACTGACCGTCTGGGACACAAAGAAC
 GLNTHRARGVALSERAL AILEGLULYSTYRLEULYSASPGLNALAGLNLEUASNALATRP
 TCCAGACTAGGGTCTGCCATCGAGAAGTACTTAAAGGACCAGGGCAGCTAAATGCTT
 7900
 GLYCYSALAPHEARGGLNVALCYSHISTRH RVALPROTRP PROASN ALASERLEUTHR
 GGGGATGTGGCTTAGACAAGTCTGTCAACTACTGTACCATGCCAAATGCAAGTCTAA
 PROASPTRPASN ASNGLUTHRTRPGLNGLUTRPGLUARGLYSVALASP PHELEUGLUALA
 CACCAGATTGGAAACAATGAGACTTGGCAAGAGTGGAGCCAGGTTGACTTCTGGAGG
 8000
 ASNILETHR ALA EULEU GLUGLU ALAGLNILEGLNGLULYSASN M ETTYRGLULEU
 CAAATATAACGGCCCTCCTAGAACAGGGCACAATTCAACAAAGAGAACATGTATGAAT
 8100
 GLNLYSLEU ASN SERTRP ASPVALPHEGLYASNTRP PHEASPLEU THR SERTRPILEY
 TACAAAAGTTGAATAGCTGGATGTGTTGGCAATTGGTTGACCTTACTTCTGGATAA
 TYRILEGLNTYRGLYILETYRILEILEVALGLYVALILELEU ARGILEVALILETYR
 AGTATATAACATATGGAATTATATAATTGTAGGAGTAATCTGTTAACAGTGTATGATCT
 8200
 ILEVALGLNMETLEU ALA ARGLEU ARGGLNGLYTYRARGPROVALPHE SER SER PRO
 ATAGTACAAATGCTAGCTAGCTAACAGGGTATAGGCCAGTGTCTTCCCCAC
 TAT2ARGPROILEPROASNARGILEARGLEUCYSGLNPROLYSLYSALA
 ART2VALASPPROT YRPROTHR GLYSER GLYSER ALA ASN GLNARG ARG GLN
 SERTYR PHE GLN***THR HISTHR GLN GLN ASPPRO ALA LEU PROTHR GLY SGLUG LYLYS
 CCTCTTATTCCAGTAGACCCATACCCAACAGGATCCGGCTCTGCCAACCAAAGAAC
 8300
 LYSLYSGLUTHRVALGLUA ALAVAL ALA LATHRALA PROGLY LEUGLY ARG TAT (fin)
 LYSARGARGTRPARGGLNARGTRPGLNGLNLEULEU ALA LEU ALA SPARGILETYR
 LYSGLYASPGLYGLYGLYSERGLYGLYASN SER SER TRP PROTRP GLN ILE GLU TYRILE
 AAAAAGGAGACGGTGGAGGCAGCGGTGGCAACAGCTCCTGGCCTTGGCAGATAGAATATA
 8400

(fig. 1B-suite 6)

17/35

SERPHEPROASPPROPROTHRASPTHRPROLEUASPLEUALAILEGLNGLNL^EUGLN^NS^N
 HISPHELEUILEARGLNL^EU^ILEARGL^EULEUTHRTRPLEUPHESERASNCYSARGTHR
 TTCA^TTTCC^TGATCCGCCAACTGATA^CGGCTTGA^TGGCTATT^CAGCAACTGCCAGAA
 LEUALAILEGLUSERILEPROASPPROPROTHRASNILEPROGLUALALEUCYSA^PLFU
 LEULEUSERARGALATYRGLNL^ILEU^EUGLN^PROILEPHEGLNARGLEUSERALATHRTYR
 CCTTGCTATCGAGAGCATA^CACAGATCCTCCAACCAATATTCCAGGGCTCTGGGACCT
 8500 F METGLYGLYALA
 ARGARGILEARGARGSERPROGLNALA ART2 (fin)
 GLYGLUPHEGLYGLUVALLEUARGLEUGLULEUTHRTYRLEUGLNTYRCLYTRPSERTYR
 ACGGAGAAATT^CGGAGAAGTCCTCAGGCTTGA^ACTGACCTACCTACAATATGGTGGAGCT
 ILESERLYSLYSARGSERLYSPROPROGLUILECYSASPARGASP^SERCYSGLYARGVAL
 PHEGLNGLUALAYALGLN^AALAALAARGASPLEUARGGLNARGLEULEUARGALAARGGLY
 ATTTCCAAGAACGGGTCC^AGGCCAGAGATCTGCGA^CAGAGACTCTTGGGGCGCTG
 8600 GLYARGASNTYRCLYARGLEUPHELYSGLYALGLUAS^PGLYSERSERGLNSERLEUGLY
 GLULYSLEUTRPGLUALALEU^EUGLN^NARGGLYGLYARGTRPILELEUALAILEPROARGARG
 GGGAGAAATTATGGGAGGCTTCAAAGGGTGGAA^GATGGATCCTCGCAATCCC^TAGGA
 8700 GLYLEUASPLYSGLYLEUSERSERLEUSERS^CYSGLU^LYGLNLYSTYRA^SNGLNGLYGLU
 ILEARGGLNGLY^ILEUGL^LULEUTHRLEULEU
 GGATTAGACAAGGGCTTGAGCTACTCTGTGAGGGCCAAAAATACAATCAGGGAGAA
 TYRHETASNTHRPROTRPARGASNPROALAGLU^GLYSLY^SLEUPROTYRARGLYS
 TACATGAATACTCCATGGAGAAACC^CAGCTGAAGAGAGGAAAAATTACCATACAGAAAA
 8800 GLNASHNILEASPASPILEASPGLUGLUASPASPASPLEUVALGLYILEPROVALGLUALA
 CAAAATATAGATCATATAGATGAGGAAGATGATGACTTGGTAGGGATACCAGTTGAGGCC
 ARGYLPROLEUARGTHR^MET^TYR^RLYSLEUALAILEASP^METSERHIS^PH^EILELYS
 AGAGTTCCC^TAAGAACAA^TGAGTTACAAATTGGCAATAGATATGTCTCATTTATAAAA
 8900 GLULYSGLYGLY^ILEUGL^LUGLYILEILEPROASPTRPGLNILEHIS^SERGLYPROGLYILE
 GAAAAGGGGGACTGGAA^GGGATTATTACAGTGCAAGAACATAGAA^TCTTAGACATA
 9000 TYRLEUGLULYSGLUGLUGLYILEILEPROASPTRPGLNILEHIS^SERGLYPROGLYILE
 TACTTAGAAAAGGAAGAAGGCATCAC^AGGATTGGCAGATA^CACTCCGGACCAGGAATT
 ARGTYREULYSMETPHEGLYTRPLEUTRPLYSLEU^ILEPROVALASNVALSERAS^PGLU
 AGATACCTAAAGATGTTGGCTGGCTATGGAAATTAA^TCCCTGTAATGTATCAGATGAG
 9100 ALAGLNGLUAS^PGLUGLUHIS^TYRLEUVALHIS^SPROALAGLN^HR^SERGLN^TRPASPASP
 GCACAGGAGGATGAGGAGCATTATTAGTGCA^CCCAGCTCAAAC^TCCCAGTGGGATGAC
 PROTRPGLYGLUVALLEUALATRPLYS^HASP^PPROTHRLEUALATYRTHRTYRGLUALA
 CCTTGGGGAGAGGTTCTAGCATGGAA^GTTGATCCA^ACTCTAGCCTACACTTATGAGGCA
 9200 TYRILEARGTYRPROGLUGLUPHEGLYSERLYSSERGLY^ILEUSERGLULYSGLUVALLYS
 TATATTAGATACCAGAAGAGTTGGAA^GCAAGTCAGGCTGT^CAGAGAAAGAGGTTAAA
 9300 ARGARGLEUALAALAARGGLY^ILEULEUGLUMETALAASP^PARGLY^SGLUTHRSER
 AGAAGGCTAGCC^GCAAGAGGCC^TCTTGA^AATGGCTGACAGGAAGGAAACTAGCTGAGAC
 ACCAGGGACTTTCCACAAAGGGATGT^CATGGGGAGGTACTGGGGAGGAGCCGGTTGGAA
 9400 CACCCACTTTCTGATGATAAATATCAC^TGCATT^TCGCTCTGTATT^CAGTCGCTCTGCG
 GAGAGGCTGGCAGATTGAGCC^TGGAGGTCTCTCCAGCACTAGCAGGTAGAGCCTGG
 9500 TGTTCCCTGCTAGACTCTACCAGCACTGGCCGGTGTGGCAGAGTGGCTCCACGCTT
 9600

(fig.1B-suite 7)

18 / 35

FIG. 1C

séquence LTR
CIVET
versus
HIV-2 ROD

X 8960 8970 8980 8990 9000 9010
TCGAAGGGATTATTACAGTGCAAGAACATAGAATCTTAGACATATACTTAGAAAAAGG
::: :::: : ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
TGGAAAGGGATGTTTACAGTGAAAGAACATAAAATCTTAAATATACTTAGAAAAAGG
X 8950 8960 8970 8980 8990

9020 9030 9040 9050 9060
AAGAAGGCATCATACCAGATTGGCAGATACACTCCGGA---CCAGGAATTAGATACCTAA
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
AAGAAGGGATAATTGCAGATTGGCAGAACTACACTCATGGGCCAGGAGTAAGATAACCAA
9010 9020 9030 9040 9050

9080 9090 9100 9110 9120
AGATTTGGCTGGCTATGGAAATTAAATCCCTGTAAATGTATCAGATGAGGCACAGGAGG
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
TGTTCTTGGGTGGCTATGGAAAGCTAGTACCAAGTAGATGTCCCACAAGAAGGGGAGGGACA
9070 9080 9090 9100 9110

9140 9150 9160 9170 9180
ATGAGGAGCATTATTAGTGCACCCAGCTCAAACCTCCCAGTGGGATGACCCCTGGGGAG
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
CTGAGACTCACTGCTTAGTACATCCAGCACAAACAAGCAAGTTGATGACCCGCATGGGG
9130 9140 9150 9160 9170

9200 9210 9220 9230 9240
AGGTTCTAGCATGGAAAGTTGATCCAACCTCTAGCCTACACTTATGAGGCATATATTAGAT
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
AGACACTAGTCTGGAGTTGATCCCTTGCTGGCTTATAGTTACGAGGGCTTTATTGGT
9190 9200 9210 9220 9230

9260 9270 9280 9290 9300
ACCCAGAAGAGTTGGAAGCAAGTCAGGCCTGTCAGAGAAAGAGGTTAAAAGAAGGCTAG
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
ACCCAGAGGAATTGGGCACAAGTCAGGCCTGCCAGAGGAAGAGTGGAAAGGGCGAGACTGA
9250 9260 9270 9280 9290

9320 9330 9340 9350
CCGCAAGAGGCCTTCTGAAATGGCT-GACAGGAAGGAAACT-----
::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
AAGCAAGAGGAATACCATTAGTTAAAGACAGGAACAGCTATACTTGGTCAGGGCAGGAA
9310 9320 9330 9340 9350

FIG. 1C

19 / 35

(fig. 1C-suite 1)

20/35

FIG. 2
 (HIV-2.P
 versus
 (HIV-1.P

..... env4

HIV2----- MIGQLLIA ILLA-SACLV YCTCYVTVFV GVPTIUKNATI
 HIV1----- HRVKEKYQHL WRNGWKGTM LLGILMICSA TEKLWVTVYY GVPVWKEATT

5

..... env5

HIV2----- PLFCATRNR- DT WG TIQCLPDNDD YOEITL-NVT EAFDAWNNTV
 HIV1----- TLFCASDAKA YDTEVBNVVA THACVPTDPN PQEVVLVNVT ENFNNWKNDM

.....

10

110 env6 120 130 140 150
 HIV2----- TEQAIEDVWH LFETSIKPCV KLTPLCVAMKICSSSTESSTGN NTTSKSTSTT
 HIV1----- VEQMHEDIIS LWDQSLKPCV KLTPLCVSLK CTDL---CH ATNTHSSNTN

.....

15

160 170 180 190 200
 HIV2----- -TTPTDQE QEISEDTCA RADNCGLGE SETINCQFMN TGLERDKKKQ
 HIV1----- SSSGENMHEK GEIK --- HCSFNIS TSIRGKVQKE YAFFYKLDII

.....

20

210 220 230 env7 240 250
 HIV2----- Y--NET-WYS KVYCETNNST NQTQCYMNNHC NTSVITESCD KHYHDAIREE
 HIV1----- PIDNDTTSYT ----- TSC NTSVITQACP KVSFEPIIH

.....

25

260 env8 270 280 290 300
 HIV2----- YCAPPGYALL RC-NDT-NYS GFAPNCSKVV ASTCTRMMET QTSTWF-GFN
 HIV1----- YCAPAGFAIL KCNNKTENGTP GP---CTNVS TVQGTHGIRP VVSTQLLL-N

.....

30

310 320 330 340 350
 HIV2----- GTRAE---H RTIYIYWHGRD N-RTII-SLN KYYNLSLCK RPGNKTVKQI
 HIV1----- CSLAEEEVVI RSAFT---D NAKTIIVQLN QSVE--INCT RPNNNTRKSI

.....

360 370 env9 380 390 400
 HIV2----- NLNS--GHVF HSKYQPINKR PROAWCWFKG -RKWDANQEV KETLAKHPRY
 HIV1----- RIQRGPGEAF VTIGKIGN-- MRQAHCHISR AKWIDAT---L KQIASKLREQ

.....

FIG. 2

21 / 35

| | | | | | | | | |
|----|-----------|---|--|--|--|--|-------|-------|
| | | 410 | ↓ | 420 | env10 | 430 | 440 | 450 |
| | HIV2----- | RGTNDTRNIS | FAAPGKGS <u>D</u> P | EVAYMW <u>T</u> NCR | GEFLYCNNTW | FLH--WI--- | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | FGNNKT--II | FKQSS-GGD <u>P</u> | EIVTHSFNC <u>G</u> | GEFFYCNST <u>Q</u> | LFNSTWFNST | | |
| | | | | | | | | |
| 5 | HIV2----- | 460 | ↓ | 470 | env11 | 480 | 490 | 500 |
| | | EN | KTHRNYAP <u>C</u> H | I <u>K</u> OIINTW <u>H</u> R | VGRNVY <u>L</u> PPR | E <u>G</u> E <u>L</u> SCN <u>S</u> T <u>V</u> | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | WSTE <u>G</u> SNNT <u>E</u> | GS <u>D</u> TITL <u>P</u> CR | I <u>X</u> Q <u>F</u> INMW <u>Q</u> E | V <u>G</u> KAM <u>Y</u> APP <u>I</u> | SG <u>Q</u> IR <u>C</u> SS <u>N</u> I | | |
| | | | | | | | | |
| 10 | HIV2----- | 510 | 520 | 530 | 540 | 550 | | |
| | | TSIIANIDW <u>Q</u> | NNNQT <u>N</u> ITFS | A <u>E</u> V <u>A</u> E <u>L</u> Y <u>R</u> L <u>-</u> | —EL <u>G</u> D <u>Y</u> K <u>L</u> V | EITPI <u>G</u> FAPT | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | T <u>G</u> LLL <u>T</u> R <u>D</u> GG | NNNN <u>G</u> SEIFR | P <u>G</u> CCD <u>M</u> HRDN <u>W</u> | R <u>S</u> ELY <u>K</u> Y <u>K</u> V <u>V</u> | KIEPLGVAPT | | |
| | | | | | | | | |
| 15 | env3 | 560 | 570 | 580 | 590 | 600 | | |
| | HIV2----- | <u>K</u> E <u>K</u> R <u>Y</u> <u>S</u> <u>S</u> <u>A</u> <u>H</u> <u>G</u> | R <u>H</u> TR <u>G</u> V <u>F</u> V <u>L</u> <u>G</u> | —FL <u>G</u> FL <u>A</u> TA | GS <u>A</u> M <u>G</u> A <u>A</u> S— | LTV <u>S</u> A <u>Q</u> S <u>R</u> T <u>L</u> | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | KAKRR--VV <u>Q</u> | REKRAVGI-G | ALFL <u>G</u> FL <u>G</u> AA | G <u>S</u> T <u>M</u> G <u>A</u> R <u>S</u> MT | LTVQA--RQL | | |
| | | | | | | | | |
| 20 | HIV2----- | 610 | 620 | 630 | ↓ 640 | env1 | 650 | |
| | | LAGIVQQQQQ | LLDVVKR <u>Q</u> QE | LLRLTVNGTK | N <u>L</u> Q <u>A</u> R <u>V</u> T <u>A</u> I <u>E</u> | KYLO <u>D</u> Q <u>A</u> R <u>L</u> N | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | LS <u>G</u> IVQQQN <u>N</u> | LLRAI <u>E</u> A <u>Q</u> Q <u>H</u> | LLQLTVWGIK | QL <u>Q</u> AR <u>I</u> L <u>A</u> V <u>E</u> | RYLK <u>D</u> Q <u>Q</u> LL <u>G</u> | | |
| | | | | | | | | |
| 25 | HIV2----- | 660 | 670 | 680 | 690 | 700 | | |
| | | <u>S</u> WG <u>C</u> A <u>F</u> R <u>Q</u> <u>V</u> <u>C</u> | <u>H</u> T <u>T</u> V <u>P</u> <u>W</u> <u>—</u> | V <u>N</u> D <u>S</u> L <u>A</u> P <u>D</u> W <u>D</u> | N <u>M</u> T <u>W</u> Q <u>E</u> W <u>E</u> K <u>Q</u> | V <u>R</u> Y <u>L</u> E <u>A</u> N <u>I</u> S <u>K</u> | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | I <u>W</u> GC <u>S</u> G <u>K</u> L <u>I</u> C | TTAV <u>Z</u> W <u>H</u> A <u>S</u> I. | SN <u>K</u> S <u>L</u> E <u>Q</u> I <u>W</u> U | N <u>H</u> T <u>W</u> W <u>E</u> D <u>R</u> E | I <u>N</u> NY <u>T</u> S <u>L</u> I <u>N</u> S | | |
| | | | | | | | | |
| 30 | HIV2----- | 710 | 720 | 730 | 740 | 750 | | |
| | | <u>S</u> LE <u>Q</u> A <u>Q</u> I <u>Q</u> Q <u>E</u> | <u>K</u> N <u>M</u> Y <u>E</u> L <u>Q</u> K <u>L</u> N | <u>S</u> W <u>D</u> I <u>F</u> G <u>N</u> W <u>F</u> D | L <u>T</u> SW <u>V</u> K <u>Y</u> I <u>Q</u> Y | G <u>V</u> L <u>I</u> I <u>V</u> A <u>I</u> A | | |
| | | * | * | * | * | * | * | * |
| | HIV1----- | L <u>I</u> E <u>E</u> S <u>Q</u> N <u>Q</u> Q <u>E</u> | K <u>N</u> E <u>Q</u> E <u>L</u> L <u>E</u> L <u>D</u> | K <u>W</u> A <u>S</u> L <u>W</u> N <u>W</u> F <u>H</u> | I <u>T</u> N <u>W</u> L <u>W</u> Y <u>I</u> K <u>I</u> | F <u>I</u> M <u>I</u> V <u>G</u> G <u>L</u> V <u>G</u> | | |
| | | | | | | | | |

(fig.2 - suite 1)

22 / 35

| | 760 | 770 | 780 | 790 | 800 |
|-----------|------------------|------------|--------------|--------------|-----------------|
| HIV2----- | LRIVIYVVQM | LSRLRKGYRP | V-FSSPPGYI | QQIHIHKDRG | QPANEETEED |
| | *** * * * * | | * * | | ** * ** |
| HIV1----- | LRIVFAVLSI | VNRVRQGYSP | LSFQT----- | -----ULPTPRG | PDRPEGIEEE |
| ***** | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| | 810 | 820 | 830 | 840 | 850 |
| HIV2----- | GGSNGCDRYW | PWPIAYIRFL | IRQLIRLLT- | -----LYSIC | RDLLSRSFLT |
| | ** ** | | * * * | | **** |
| HIV1----- | GGERDRDRSI | RLVNGSLA-L | IWDDLRSLCL | FSYHRL----- | RDLLLIVTRI |
| ***** | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| | 860 | 870 | 880 | 890 | 900 |
| HIV2----- | LQLIYQNL RD | WLRLRTA-F | LQYGCEWIQE | AFQ-----AAA | RATRETL--- |
| | * * * | | *** | | * * * |
| HIV1----- | VELLG--RRG | WEALKYWWNL | LQYWSQELKN | SAVSLLNATA | IAVAEGTDRV |
| ***** | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| | 910 | 920 | 930 | 938 | |
| HIV2----- | ----AGACRG | LWRVLERIGR | GILAVP RRI R | QGA E I ALL | |
| | ** | | ***** | ** * ** | |
| HIV1----- | IEVVQGACRA | ----- | -IRHIP RRI R | QGLERILL | |
| ***** | ***** | ***** | ***** | ***** | |

(fig. 2 - suite 2)

23 / 35

FIG. 3 (ENV-mac
 (versus
 (ENV-ROD

| | | | | |
|--|-----|-----|-----|-----|
| 10 | 20 | 30 | 40 | 50 |
| MGCLGNOLLIAIC--SKCLWIICIQYYTIFYGVPARNATIPLFCATKNRDTWGTQCL ::: :::::: : :: : ::::::: :: : ::::::: :: ::::::: :: :: :: | | | | |
| MM---NOLLIAILLASACLVY-CTQYYTIFYGYPTWKNATIPLFCATRNRTDHWGTIQCL 10 20 30 40 50 | | | | |
| | | | | |
| 60 | 70 | 80 | 90 | 100 |
| PDNDDYSELALNYTESFOAHENTVTEQAIEDVWQLFETSIKPCVKLSPLCITMRCNKSET :::::: : ::::: ::::: ::::::::::::: ::::::::::::: :: :: :: :: | | | | |
| PDNDDYQEITLNVTEAFDAHNNTVTEQAIEDVWHLFETSIKPCVKLTPLCVAMKCSSTES 60 70 80 90 100 110 | | | | |
| | | | | |
| 120 | 130 | 140 | 150 | 160 |
| DKHGLTKSSTTASTTTTAKSYETROIVNETS---PCVVHDNCTGLEQEPMISCKENM :: : : ::::::: | | | | |
| STGNNTTSKST--STTTTTP----T-DQEQEISEDTPCARADNC SGLGEEETINCQENM 120 130 140 150 160 | | | | |
| | | | | |
| 180 | 190 | 200 | 210 | 220 |
| TGLKRDKKKEYNETWYSADLVCEQGNSTGNESRCYMNHCNTSVIQECCDKDYWDAIRCRY ::: ::::: ::::::: : :: : :: : ::::::::::::: : :: :: :: :: :: | | | | |
| TGLERDKKKQYNETWYSKDVYCETNNST-NOTACYMNHCNTSVITESCDKHYWDAIRFRY 170 180 190 200 210 220 | | | | |
| | | | | |
| 240 | 250 | 260 | 270 | 280 |
| CAPPGYALLRCNDTNYSGFMPNC SKVVVSCTRM METQTSTWFRFNGTRAENRTYIYWHG ::::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | |
| CAPPGYALLRCNDTNYSGFAPNC SKVVA STCTRM METQTSTWFGFNGTRAENRTYIYWHG 230 240 250 260 270 280 | | | | |
| | | | | |
| 300 | 310 | 320 | 330 | 340 |
| RDNR TIISLNKHYNLTMKCRRPGNKT VLPVTIMSALVFHS---QPVNERPKQAWCRFGGNW ::::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | |
| RDNR TIISLNKYYNLSLHCKRPGNKT VKAIMLMSGHVFHSHYQPI NKRPRQAWCHFKGK W 290 300 310 320 330 340 | | | | |
| | | | | |
| 360 | 370 | 380 | 390 | 400 |
| KEAIKEVKQTIVKHPRYTGTNN TDKINLTAPRG DPEVTFMHTNCRGEFLYCKMNWFLN ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | |
| KDAHQEVKETLAHKPRYRG TDTRNISFAAPGKGSDPEVAYWHTNCRGEFLYCNHTWFLN 350 360 370 380 390 400 | | | | |

FIG. 3

24/35

| | | | | |
|---|-----|-----|-----|---------|
| 420 | 430 | 440 | 450 | 460 |
| WVEDRSLTQKPKERHKRNYVPCHIRQIINTWHKVGVKNVYLPPREGDLTCNSTVTSIAN | | | | |
| : | | | | |
| WIEN-----KT-H-RNYAPCHIKQIINTWHKVGRNVYLPPREGELSCNSTVTSIAN | | | | |
| 410 | 420 | 430 | 440 | 450 |
| * 480 490 500 510 520 | | | | |
| INWTDGNQTSITMSAEVAELYRLELGOKLVEITPIGLAPTNVKRYTTG-GTSRNKRGVF | | | | |
| : | | | | |
| IDWONNNNOTNITFSAEVAELYRLELDYKLVEITPIGFAPTKEKRYSSAHG--RHTRGVF | | | | |
| 460 | 470 | 480 | 490 | 500 510 |
| 540 550 560 570 580 | | | | |
| VLGFLGFLATAGSAMGAASLTVAOSRTLLAGIVQQQQQLLOVVKRQQELLRLTVHGTKN | | | | |
| : | | | | |
| VLGFLGFLATAGSAMGAASLTVAOSRTLLAGIVQQQQQLLOVVKRQQELLRLTVHGTKN | | | | |
| 520 | 530 | 540 | 550 | 560 570 |
| 600 610 620 630 640 | | | | |
| LQTRYSAIEKYLQDQNAQLNMYGCAFQYCHTTYPWPNASLTPDNNTWQEWERKVDLE | | | | |
| : | | | | |
| LQARVTAIEKYLQDQARLNNSHGCAFQYCHTTYPWNDSLAPDWDNMNTWQEWEKQVYLE | | | | |
| 580 | 590 | 600 | 610 | 620 630 |
| 660 670 680 690 700 | | | | |
| ANITALLEEAQIQQEKNMYELOKLNSHDVFGNWFDLTSWIKYI0YGIYIIVGVILLRIVI | | | | |
| : | | | | |
| ANISKSLEAQIQQEKNMYELOKLNSHDIFGNWFDLTSWVKYI0YGVLIIVAVIALRIVI | | | | |
| 640 | 650 | 660 | 670 | 680 690 |
| 720 730 740 750 760 | | | | |
| YIVQMLARLRQGYRPYFSSPPSYFO*THTQQDPALPTKEGKKGDGGSGGNSSWPHQIEY | | | | |
| : | | | | |
| YVVQMLSRLRKGYRPYFSSPPGYI0QIHIHKDRGQPANEETEEDGGSNGGDRYWPWPIAY | | | | |
| 700 | 710 | 720 | 730 | 740 750 |
| 780 790 800 810 820 | | | | |
| IHFLIRQLIRLLTHLFSNCRTLLSRAYQILOPIFORLSATYGEFGEVRLRELTYLQYGWS | | | | |
| : | | | | |
| IHFLIRQLIRLLTRLYSICRDLLSRSFLLQLIYONLRDW-----LRLRTAFLQYGCE | | | | |
| 760 | 770 | 780 | 790 | 800 |
| 840 850 860 870 880 | | | | |
| YFOEAQAA-RDLRORLLRA-RGEKLHEALQRGGRWILAIPRRIRQGLELTL | | | | |
| : | | | | |
| WIQEAFQAAARATRETLAGACRG--LWRVLERIGRGILAVPRRIROGAEIALL | | | | |
| 810 | 820 | 830 | 840 | 850 |

(fig. 3-suite 1)

25/35

FIG. 4 (versus
(GAG-mac
(GAG-ROD

| | | | | | |
|---|-----|-----|-----|-----|-----|
| 10 | 20 | 30 | 40 | 50 | |
| YQHKKEIAVFPGRDNKIEHEMGARNSVLSGKKADELEKIRLRPGGKKYMLKHVVWAAN | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| MGARNSVLRGKKADELERIRLRPGGKKYRLKHIVWAAN | | | | | |
| 10 | 20 | 30 | | | |
| 70 | 80 | 90 | 100 | 110 | |
| ELDRFGLAESLLENKEGCOKILSVLAPLVPTGSENLKSLYNTVCVIWCIAEEKVKHTEE | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| KLDRCGLAESLLESKEGCOKILTVLDPMVPTGSENLKSLFNTVCVIWCIAEEKVKDTEG | | | | | |
| 40 | 50 | 60 | 70 | 80 | 90 |
| 130 | 140 | 150 | 160 | 170 | |
| AKQIVORHLVMETGTAETMPKTSRPTAPPFSRGGNYPVQQIGGNYTHLPLSPRTLNAWVK | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| AKQIVRRHLVAETGTAEKMPSTSRTAPSSEKGGNYPVQHVGGNYTHIPLSPRTLNAWVK | | | | | |
| 100 | 110 | 120 | 130 | 140 | 150 |
| 190 | 200 | 210 | 220 | 230 | |
| LIEEKKFGAEVYSGFOALSEGCLPYDINQMLNCVGDHQAMQIIRDIIINEAADWDLQHP | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| LYEEKKFGAEVYVPGFOALSEGCTPYDINQMLNCVGDHQAMQIREIINEEAAEDVQHP | | | | | |
| 160 | 170 | 180 | 190 | 200 | 210 |
| 250 | 260 | 270 | 280 | 290 | |
| QQAPQQ-GQLREPSGSDIAGTTSTYEEQIQWMYRQQNPIPVGNIYRRWIQLGLQKCYRMY | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| IPGPLPAGQLREPRGSDIAGTTSTYEEQIQWMFRPQNPPVYGNIYRRWIQIGLQKCYRMY | | | | | |
| 220 | 230 | 240 | 250 | 260 | 270 |
| 300 | 310 | 320 | 330 | 340 | 350 |
| NPTNILDVKQGPKEPFQSYVDRFYKSLRAEQTOPAVKNWMTOTLLIONANPDCKLVLKGL | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| NPTNILDIKQGPKEPFQSYVDRFYKSLRAEQTOPAVKNWMTOTLLVQANANPDCKLVLKGL | | | | | |
| 280 | 290 | 300 | 310 | 320 | 330 |
| 360 | 370 | 380 | 390 | 400 | 410 |
| GTNPTLEEMLTACQGVGGPGQKARLMAEALKEALAPAPIFAAAQQKGPRKPIKCHNCGK | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| GMNPTLEEMLTACQGVGGPGQKARLMAEALKEVIGPAPIFAAAQQ---RKAFCCHNCGK | | | | | |
| 340 | 350 | 360 | 370 | 380 | 390 |

FIG. 4

26 / 35

(fig.4 - suite 1)

27 / 35

(POL-mac
(versus
(POL-ROD

FIG. 5

28/35

| | | | | | |
|---|------|------|------|------|-----|
| 410 | 420 | 430 | 440 | 450 | 460 |
| DRYVLQLKELLNSIGFSSPEEKFKOPPFQWMGYELWPTKWKLOKIELPORETHTYNDIQ | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| DRYVLQLKELLNGLGSTPDEKFQKOPPYHWMGYELWPTKWKLOKIQLPOKEINTVNDIQ | | | | | |
| 390 | 400 | 410 | 420 | 430 | 440 |
| 470 | 480 | 490 | 500 | 510 | 520 |
| KLYGVLNHAAQIYPGIKTKHLTRLIRGKMTLEEVOWTEMAEAEEENKIILSQEQQEGCY | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| KLYGVLNHAOLYPCIKTKHLTRLIRGKMTLEEVOWTELAEAELEENRITLSQEQQEGHY | | | | | |
| 450 | 460 | 470 | 480 | 490 | 500 |
| 530 | 540 | 550 | 560 | 570 | 580 |
| YQESKPLEATVIKSODNQWSYKIHOEDKILKVGKFAIKNTHTNGVRLLAHVIDKICKEA | | | | | |
| ::: ::::: ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| YQEEKELEATYQKDQENQWTYKIHQEEKILKVGKYAKVKNTHTNGIRLLAQVVQKICKEA | | | | | |
| 510 | 520 | 530 | 540 | 550 | 560 |
| 590 | 600 | 610 | 620 | 630 | 640 |
| IVIWGQVPKFHLPLYEKDYWEQWHDYHQYTHIPEWDFISTPPLYRLYFNLYKDPIEGET | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| LVJWGRIPKFHLPLYEREIWEQWHDNYHQYTNPWDFTSTPPLYRLAFNLVGDPIPGAE | | | | | |
| 570 | 580 | 590 | 600 | 610 | 620 |
| 650 | 660 | 670 | 680 | 690 | 700 |
| YYVDGSCSKOSKEGKAGYITDRGKDKVKVLEOTTNQAELEAFALHALTDSGPKANIIVDS | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| FYTOGSCNRQSKEGKAGYYTDRGKDKVKVLEOTTNQAELEAFALHALTDSGPKVNIIVDS | | | | | |
| 630 | 640 | 650 | 660 | 670 | 680 |
| 710 | 720 | 730 | 740 | 750 | 760 |
| QYVMGIIITGCPTESESRLVNOIIIEHIKTEIYYVANVPAHKGIGGNQEIDHLYSQGIRQV | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| QYVMGISAQPTESESKIVNQIIIEHIKKEAIYYVANVPAHKGIGGNQEVDHLYSQGIRQV | | | | | |
| 690 | 700 | 710 | 720 | 730 | 740 |
| 770 | 780 | 790 | 800 | 810 | 820 |
| LFLEKIEPAQEEHSKYHSNIKELVFKFGLPRLVAKQIVDTCDOKCHOKGEAIGHGQVNSOLG | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| LFLEKIEPAQEEHEKYHSVKELSHKFGIPNLVARQIVNSCAOCQQKGEAIGHGQVNAELG | | | | | |
| 750 | 760 | 770 | 780 | 790 | 800 |
| 830 | 840 | 850 | 860 | 870 | 880 |
| THQMDCTHLECKIVIVAYVHYASGFIEAEYIPOETGROTAFLKLASRWPITHLHTDNGA | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| THQMDCTHLECKIIIIVAYVHYASGFIEAEVIPQESGROTAFLKLASRWPITHLHTDNGA | | | | | |
| 810 | 820 | 830 | 840 | 850 | 860 |
| 890 | 900 | 910 | 920 | 930 | 940 |
| NFASQEYKMHAWGIEHTFGVYNPQSQGVVEAHNNHHLKNQIDRIREQANSVETIVLMA | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| NFTSQEYKMYAHWIGIEQSFGVYNPQSQGVVEAMNNHHLKNQISRIREQANTIETIVLMA | | | | | |
| 870 | 880 | 890 | 900 | 910 | 920 |
| 950 | 960 | 970 | 980 | 990 | |
| VHCMNFKRRGGIGDMTPAERLINMITTEQEJOFOQSKFKNFRVYYREGRDOLHKGPG | | | | | |
| ::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | | |
| IHCMNFKRRGGIGDMTPSERLINMITTEQEJQFLQAKNSKLDFRYYREGRDOLHKGPG | | | | | |
| 930 | 940 | 950 | 960 | 970 | 980 |
| 1010 | 1020 | 1030 | 1040 | 1050 | |
| ELLHKGEGAVILKVGTDIKVYPRRKAKIIKOYGGCKEMOSSHMEDTGEAREVA | | | | | |

(fig.5-suite 1)

29/35

::::::: ::::::: ::::::: ::::: :: : : :
SELLRKGE[GAYLVKYGTDIKIIPRRKAKIIRDYGGRQEMDSGSHLEGAREDGEHA
990 1000 1010 1020 1030

(fig. 5-suite 2)

30/35

FIG. 6 (Q.mac
 (versus
 (Q.ROD

| | | | | |
|---|-----|-----|-----|-----|
| 10 | 20 | 30 | 40 | 50 |
| MEEEKRHIVVPTWRIPERLERWHSLIKYLKYTKDLQXACYYPHHKVGWAHTCSRVIFF | | | | |
| ::: :::::::::: : : : ::::: :::::::::::: : ::::::::::::: ::::::::::::: | | | | |
| MEEDKRHIVVPTWRVPGRMEKWHSLVKYLKYTKDLEKVCYYPHHKGWAHTCSRVIFF | | | | |
| 10 | 20 | 30 | 40 | 50 |
| | | | | |
| 70 | 80 | 90 | 100 | 110 |
| LQEGSHLEYDGYHNLTPERGWLSTYAVRITWYSKDFWTDTPEYADILLHSTYFPCFTAG | | | | |
| : ::::: : ::::::: ::::: : ::::::: ::::::: :: : ::::::::::::: | | | | |
| LKGNNSHLEIQAYNLTPEKGWLSYSVRITWYTEKFWTDTPECADVLIHSTYFPCFTAG | | | | |
| 70 | 80 | 90 | 100 | 110 |
| | | | | |
| 130 | 140 | 150 | 160 | 170 |
| EVRAIRGERLLSCCRFPRAHKHQVPSLQYLALRYVSHY-RSQGENPTHKQWRDNRRSL | | | | |
| ::::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: | | | | |
| EVRAIRGEKLLSCCNYPRAHRAQVPSLQFLALVVVQQNORPQRDSTTRKQRRDYRRGL | | | | |
| 130 | 140 | 150 | 160 | 170 |
| | | | | |
| 180 | 190 | 200 | 210 | |
| RYAKQNSRGDKORGKPPTEGANFPGLAKVLGILA | | | | |
| : ::: :: : :: : :: : :: : :: : :: : | | | | |
| RLAKQDSRSRKQRSSESPTPRTYFPGVAEYLEILA | | | | |
| 190 | 200 | 210 | | |

FIG. 6

31/35

(R.mac
FIG. 7 (versus
 (R.ROD

10. 20 30 40 50
ME---ERPPENEGPQREPWDEKVVVEVLKELKEEALKHFOPRLLTALGNHIYNRHGOTLE
:
MAEAPTELPPVDTPLREPGDEWIIEILREIKEEALKHFOPRLLIALGKYIYTRHGOTLE
10 20 30 40 50

60 70 80 90 100
GAGELIRILQRALFIHFRSGCSHSRIGQPCCGNPLSTIPPSRSML
:
GARELIKYLQRALFTHFRAGCGHSRIGQTRGGNPLSAIPTPRNMQ
70 80 90 100

FIG. 7

32 / 35

FIG. 8 (X.mac
 (versus
 (X.ROD

| | | | | |
|--|--|--|--|--|
| 10 | 20 | 30 | 40 | 50 |
| MSDPRERI PPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVYORSWEYWHDEQ | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: |
| MTOPRETV PPGNSGEETIGEAFAWLNRTVEAINREAVNHLPRELIFQVWARSWRYWHDEQ | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: | : ::::: ::::::::::::: ::::::: ::::::: ::::::: ::::::: ::::::: ::::::: |
| 10 | 20 | 30 | 40 | 50 |
| | | | | |
| 70 | 80 | 90 | 100 | 110 |
| GMSOSYT KYRYLC IOKALFMHCKKGCRCLGEGHGAGGWRPGPPPPPPGLA | : ::::: ::::::::::::: ::::::: ::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: |
| GMSESYT KYRYLC IOKAVYMHVRKGCTCLGRGHGP GG WRPGPPPPPPGLV | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: | : ::::: ::::::::::::: ::::::: ::::: ::::: ::::: ::::::::::::: |
| 70 | 80 | 90 | 100 | 110 |

33 / 35

(F.mac
(versus
(F.ROD

FIG. 9

34/35

FIG.10 (TAT.mac
 (versus
 (TAT.ROD

10 20 30 40 50
METPLREQENSLESSNERSSYISEAAAIPESANLGEELSQLYRPLEACYNTCYCKC
::::: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : ::
METPLKAPESSLKSCNEPFSRTSEQDVATQELARQGEELSQLYRPLEFNNNSCYCKRCC
10 20 30 40 50

70 80 90 100 110
YHCQFCFLKKGLGISYEKSHRRRTPKKAKANTSSASNERP---IPNRIRLCOPKKAKKE
:::::
YHCQMCFLNKGLGICYERKGRRRTPKTKTHPSPT----POKSISTRGDSOPTKKQKK
70 80 90 100 110

120 130
TVEAAVATAPGLGR.
::::: :: :: :: ::
TVEATVETDTGPGR
120 130

FIG. 10

35 / 35

FIG.11 (ART.mac
versus
ART.ROD

| | | | | | | |
|---|----|----|----|----|-----|-----|
| 10 | 20 | 30 | 40 | 50 | | |
| MRSHTCEEELRRRLRLIHLLHOTSKYGLSHKSAAYRHLLYDPYPTGSGSANQRROKRRRW | | | | | | |
| : : : : : : : : : : : : | | | | | | |
| HNERADEEGLQRKRLIRLLHQTN-----PYPQCPGTASORRNRRRW | | | | | | |
| 10 | 20 | | 30 | 40 | | |
| | | 70 | 80 | 90 | 100 | 110 |
| RQRWQOLLALADRIYSFPDPPTDTPLDLAIQQLQLNLAIESIPDPPTNIPEALCDLRRIRR | | | | | | |
| : | | | | | | |
| KQRWQOILALADSIVTFPDPPTADSPLDQTIQHLQQLTIQELPDPPTHLPESQRLAET | | | | | | |
| 50 | 60 | 70 | 80 | 90 | 100 | |

SPQA

FIG. 11

INTERNATIONAL SEARCH REPORT

PCT/FR 88/00025

International Application No.

I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all)⁵

According to International Patent Classification (IPC) or to both National Classification and IPC
 4 C 07 K 7/10, 7/06, C 12 N 15/00, G 01 N 33/569,
 Int. Cl. : A 61 K 39/21, A 61 K 37/02

II. FIELDS SEARCHED

Minimum Documentation Searched⁷

| Classification System | Classification Symbols |
|-----------------------|--|
| Int. Cl. ⁴ | A 61 K 39/00, A 61 K 37/00, C 07 K 7/00, G 01 N 33/00, C 12 N 15/00 |

Documentation Searched other than Minimum Documentation
 to the Extent that such Documents are Included in the Fields Searched⁸

III. DOCUMENTS CONSIDERED TO BE RELEVANT⁹

| Category ¹⁰ | Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹² | Relevant to Claim No. ¹³ |
|------------------------|--|-------------------------------------|
| X, Y | WO, A, 86/02383 (PASTEUR) 24 April 1986, see pages 21-34, 55-60; claims 1-16 -- | 1-36 |
| X, Y | US, A, 4629783 (W.L. COSANT) 16 December 1986, see columns 15,16; claims 1-42 -- | 1-36 |
| X | EP, A, 0199301 (HOFFMANN-LA ROCHE & CO.) 29 October 1986, see columns 27-42; claims 1-42 -- | 1,2,6-12,14-21,24 |
| X | EP, A, 0187041 (GENENTECH) 9 July 1986, see pages 79-81, claims 1-18; pages 82,83, claims 24-38 -- | 1,2,6-12,14-21,24 |
| Y | Science, vol. 232, 1985, American Association for the Advancement of Science, (Washington, DC, US) P.J. Kanki et al.: "New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-III _{AGM}) pages 238-243 see the whole document -- | 1-36 |
| Y | Science, vol. 233, 18 July 1986, American Association for the Advancement of Science, -- | 1-36 |

¹⁰ Special categories of cited documents:¹¹ "A" document defining the general state of the art which is not considered to be of particular relevance¹² "E" earlier document but published on or after the international filing date¹³ "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)¹⁴ "O" document referring to an oral disclosure, use, exhibition or other means¹⁵ "P" document published prior to the international filing date but later than the priority date claimed¹⁶ "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention¹⁷ "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step¹⁸ "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art¹⁹ "&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

Date of Mailing of this International Search Report

10 June 1986 (10.06.86)

7 July 1988 (07.07.88)

International Searching Authority

Signature of Authorized Officer

EUROPEAN PATENT OFFICE

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

| Category* | Citation of Document, with indication, where appropriate, of the relevant passages | Relevant to Claim No |
|-----------|--|----------------------|
| | (Washington, DC, US) F. Clavel et al.: "Isolation of a new human retrovirus from West African patients with AIDS", pages 343-346 see the whole document -- | |
| Y | Nature, vol. 324, 18/25 December 1986, MacMillan Eds. (Londres, GB) F. Clavel et al.: "Molecular cloning and polymorphism of the human immune deficiency virus type 2", pages 691-695 see the whole document -- | 1-36 |
| Y | Nature, vol. 321, 22 May 1986 MacMillan Eds. (Londres, GB) M. Murphey-Corb et al.: "Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys", pages 435-437 see the whole document -- | 1-36 |
| P,X | WO, A, 87/02038 (ONCOGEN) 9 April 1987, see pages 114-117; claims 31-55 -- | 1,2,6-12,14-21,24 |
| P,X | Journal of Cellular Biochemistry, Supplement 11D, Ucla Symposia on Molecular & Cellular Biology, 29 March - 1 May 1987, Symposium on human retroviruses, 16th Annual Meeting UCLA, see page 44, abstract P112: "Human retroviruses, cancer and AIDS: approaches to prevention and therapy" -- | 1-36 |
| P,Y | FR, A, 2593189 (PASTEUR) 24 July 1987, see page 8, lines 17-26; pages 13-15, claims 1-18 -- | 1-36 |
| P,Y | Nature, vol. 326, No. 6113, 9-15 April 1987, MacMillan Eds. (Londres, GB) H. Kornfeld et al.: "Cloning of HTLV-4 and its relation to simian and human immunodeficiency viruses" pages 610-613, see the whole document -- | 1-36 |
| P,Y | Nature, vol. 326, 16 April 1987, MacMillan Eds. (Londres, GB) M. Guyader et al.: "Genome organization and transactivation of the human immunodeficiency virus type 2, pages 662-669, see the whole document -- | 1-36 |
| P,Y | FEBS Letters, vol. 218, No. 2, June 1987, Eds. Elsevier Siencc Publishers B.V. (Biomedical Division), 1987 Federation of European Biochemical Societies M.J.E. Sternberg et al.: "Prediction of antigenic determinants and secondary structures of the major AIDS virus proteins", pages 231-237 -- | 1-36 |

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

see the whole document

V. OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE¹

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. Claim numbers because they relate to subject matter not required to be searched by this Authority, namely:

2. Claim numbers, because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful International search can be carried out, specifically:

3. Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 8.4(a).

VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING²

This International Searching Authority found multiple inventions in this International application as follows:

Claims 1, 2, 6-21, 23-36

Claims 1, 2, 6-21, 23-36 all in part 3-5

Claims 23, 25-28, 34-36 all in part 22, 24, 29

1. As all required additional search fees were timely paid by the applicant, this International search report covers all searchable claims of the International application.

2. As only some of the required additional search fees were timely paid by the applicant, this International search report covers only those claims of the International application for which fees were paid, specifically claims:

3. No required additional search fees were timely paid by the applicant. Consequently, this International search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

The additional search fees were accompanied by applicant's protest.

No protest accompanied the payment of additional search fees.

**ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO.**

FR 8800025

SA 20445

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 23/06/88. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

| Patent document cited in search report | Publication date | Patent family member(s) | | Publication date |
|---|---------------------|----------------------------|----------|---------------------|
| WO-A- 8602383 | 24-04-86 | FR-A- | 2571968 | 25-04-86 |
| | | AU-A- | 5061785 | 02-05-86 |
| | | EP-A- | 0201540 | 20-11-86 |
| | | JP-T- | 62500592 | 12-03-87 |
| | | WO-A- | 8604336 | 31-07-86 |
| | | AU-A- | 5320086 | 13-08-86 |
| | | EP-A- | 0211022 | 25-02-87 |
| | | JP-T- | 62502095 | 20-08-87 |
| US-A- 4629783 | 16-12-86 | EP-A- | 0201716 | 20-11-86 |
| | | WO-A- | 8606414 | 06-11-86 |
| | | AU-A- | 5572786 | 16-10-86 |
| | | AU-A- | 5773386 | 18-11-86 |
| | | EP-A- | 0220273 | 06-05-87 |
| | | JP-T- | 62502617 | 08-10-87 |
| | | AU-B- | 571128 | 31-03-88 |
| EP-A- 0199301 | 29-10-86 | AU-A- | 5636386 | 23-10-86 |
| | | JP-A- | 62012799 | 21-01-87 |
| EP-A- 0187041 | 09-07-86 | JP-A- | 61233700 | 17-10-86 |
| WO-A- 8702038 | 09-04-87 | AU-A- | 6299286 | 09-04-87 |
| | | BE-A- | 905492 | 25-03-87 |
| | | GB-A- | 2181435 | 23-04-87 |
| | | FR-A- | 2587720 | 27-03-87 |
| | | SE-A- | 8604007 | 26-03-87 |
| | | NL-A- | 8602422 | 16-04-87 |
| | | FR-A- | 2593519 | 31-07-87 |
| | | JP-A- | 63068075 | 26-03-88 |
| FR-A- 2593189 | 24-07-87 | WO-A- | 8704459 | 30-07-87 |
| | | AU-A- | 6891187 | 14-08-87 |
| | | EP-A- | 0239425 | 30-09-87 |

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale N° PCT/FR 88/00025

I. CLASSEMENT DE L'INVENTION (si plusieurs symboles de classification sont applicables, les indiquer tous)*

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

CIB⁴ : C 07 K 7/10, 7/06, C 12 N 15/00, G 01 N 33/569,
A 61 K 39/21, A 61 K 37/02

II. DOMAINES SUR LESQUELS LA RECHERCHE A PORTÉ

Documentation minimale consultée *

| Système de classification | Symboles de classification |
|---------------------------|--|
| CIB ⁴ | A 61 K 39/00, A 61 K 37/00, C 07 K 7/00, G 01 N 33/00, C 12 N 15/00 |

Documentation consultée autre que la documentation minimale dans la mesure où de tels documents font partie des domaines sur lesquels la recherche a porté *

III. DOCUMENTS CONSIDÉRÉS COMME PERTINENTS¹⁰

| Catégorie * | Identification des documents cités, ¹¹ avec indication, si nécessaire, des passages pertinents ¹² | N° des revendications visées ¹³ |
|-------------|--|--|
| X, Y | WO, A, 86/02383 (PASTEUR) 24 avril 1986, voir pages 21-34, 55-60; revendications 1-16 -- | 1-36 |
| X, Y | US, A, 4629783 (W.L. COSANT) 16 décembre 1986, voir colonnes 15,16; revendications 1-42 -- | 1-36 |
| X | EP, A, 0199301 (HOFFMANN-LA ROCHE & CO.) 29 octobre 1986, voir colonnes 27-42; revendications 1-42 -- | 1,2,6-12, 14-21,24 |
| X | EP, A, 0187041 (GENENTECH) 9 juillet 1986, voir pages 79-81, revendications 1-18; pages 82,83, revendications 24-38 -- | 1,2,6-12, 14-21,24 |
| Y | Science, vol. 232, 1985, American Association for the Advancement of Science, (Washington, DC, US) P.J. Kanki et al.: "New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-IIIAGM)" | 1-36 |

* Catégories spéciales de documents cités: 11

- « A » document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- « E » document antérieur, mais publié à la date de dépôt international ou après cette date
- « L » document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- « O » document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- « P » document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- « T » document ultérieur publié postérieurement à la date de dépôt international ou à la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- « X » document particulièrement pertinent: l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive
- « Y » document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier.
- « & » document qui fait partie de la même famille de brevets

IV. CERTIFICATION

Date à laquelle la recherche internationale a été effectivement achevée

10 juin 1986

Date d'expédition du présent rapport de recherche internationale

- 7 JUL 1988

Administration chargée de la recherche internationale
OFFICE EUROPEEN DES BREVETS

Signature du fonctionnaire autorisé

J.C.G. VAN DER PUTTEN

| III. DOCUMENTS CONSIDÉRÉS COMME PERTINENTS | | (SUITE DES RENSEIGNEMENTS INDICUÉS SUR LA DEUXIÈME FEUILLE) |
|--|---|---|
| Catégorie* | Identification des documents cités, avec indication, si nécessaire, des passages pertinents | N° des revendications visées |
| | pages 238-243 voir le document en entier -- | |
| Y | Science, vol. 233, 18 juillet 1986, American Association for the Advancement of Science, (Washington, DC, US) F. Clavel et al.: "Isolation of a new human retrovirus from West African patients with AIDS", pages 343-346 voir le document en entier -- | 1-36 |
| Y | Nature, vol. 324, 18/25 décembre 1986, MacMillan Eds. (Londres, GB) F. Clavel et al.: "Molecular cloning and polymorphism of the human immune deficiency virus type 2", pages 691-695 voir le document en entier -- | 1-36 |
| Y | Nature, vol. 321, 22 mai 1986, MacMillan Eds. (Londres, GB) M. Murphey-Corb et al.: "Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys", pages 435-437 voir le document en entier -- | 1-36 |
| P,X | WO, A, 87/02038 (ONCOGEN) 9 avril 1987, voir apges 114-117; revendicaions 31-55 -- | 1,2,6-12, 14-21,24 |
| P,X | Journal of Cellular Biochemistry, Supplement 11D, UCLA Symposia on Molecular & Cellular Biology, 29 mars - 1 mai 1987, Symposium on human retroviruses, 16th Annual Meeting UCLA, voir page 44, abrégé P112: "Human retroviruses, cancer and AIDS: approaches to prevention and therapy" -- | 1-36 |
| P,Y | FR, A, 2593189 (PASTEUR) 24 juillet 1987, voir page 8, lignes 17-26; pages 13-15, revendications 1-18 -- | 1-36 |
| P,Y | Nature, vol. 326, no. 6113, 9-15 avril 1987, MacMillan Eds. (Londres, GB) H. Kornfeld et al.: "Cloning of HTLV-4 and its relation to simian and human immunodeficiency viruses" pages 610-613, voir le document en entier -- | 1-36 |

| III. DOCUMENTS CONSIDÉRÉS COMME PERTINENTS (SUITE DES RENSEIGNEMENTS INDICUÉS SUR LA DEUXIÈME FEUILLE) | | |
|--|---|---------------------------------|
| Catégorie * | Identification des documents cités, avec indication, si nécessaire, des passages pertinents | N° des revendications visées |
| P,Y | Nature, vol. 326, 16 avril 1987, MacMillan Eds. (Londres, GB) M. Guyader et al.: "Genome organization and transactivation of the human immuno- deficiency virus type 2, pages 662-669, voir le document en entier | 1-36 |
| P,Y | FEBS Letters, vol. 218, no. 2, juin 1987, Eds. Elsevier Siencc Publishers B.V. (Biomedical Division), 1987 Federation of European Biochemical Societies M.J.E. Sternberg et al.: "Prediction of antigenic determinants and secondary structures of the major AIDS virus proteins", pages 231-237 voir le document en entier | 1-36 |
| | ----- | |

SUITE DES RENSEIGNEMENTS INDICÉS SUR LA DEUXIÈME FEUILLE

V. OBSERVATIONS LORSQU'IL A ÉTÉ ESTIMÉ QUE CERTAINES REVENDICATIONS NE POUVAIENT PAS FAIRE L'OBJET D'UNE RECHERCHE¹

Selon l'article 17.2) a) certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

VI. OBSERVATIONS LORSQU'IL Y A ABSENCE D'UNITÉ DE L'INVENTION²

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la présente demande internationale, c'est-à-dire:

Revendications 1,2,6-21,23-36

Revendications 1,2,6-21,23-36 toutes partiellement 3-5

Revendications 23, 25-28, 34-36 toutes partiellement 22, 24, 29

1. Comme toutes les taxes additionnelles demandées ont été payées dans les délais, le présent rapport de recherche internationale couvre toutes les revendications de la demande internationale pouvant faire l'objet d'une recherche.

2. Comme seulement une partie des taxes additionnelles demandées a été payée dans les délais, le présent rapport de recherche internationale couvre seulement celles des revendications de la demande pour lesquelles les taxes ont été payées, c'est-à-dire les revendications :

3. Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale est limité à l'invention mentionnée en premier dans les revendications; elle est couverte par les revendications numéros:

4. Etant donné que toutes les revendications susceptibles de faire l'objet d'une recherche le pouvaient sans effort particulier justifiant une taxe additionnelle, l'administration chargée de la recherche internationale n'a sollicité le paiement d'aucune taxe additionnelle.

**ANNEXE AU RAPPORT DE RECHERCHE INTERNATIONALE
RELATIF A LA DEMANDE INTERNATIONALE NO.**

FR 8800025
SA 20445

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche international visé ci-dessus.

Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du 23/06/88.
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

| Document brevet cité au rapport de recherche | Date de publication | Membre(s) de la famille de brevet(s) | Date de publication |
|---|------------------------|--|--|
| WO-A- 8602383 | 24-04-86 | FR-A- 2571968 AU-A- 5061785 EP-A- 0201540 JP-T- 62500592 WO-A- 8604336 AU-A- 5320086 EP-A- 0211022 JP-T- 62502095 | 25-04-86 02-05-86 20-11-86 12-03-87 31-07-86 13-08-86 25-02-87 20-08-87 |
| US-A- 4629783 | 16-12-86 | EP-A- 0201716 WO-A- 8606414 AU-A- 5572786 AU-A- 5773386 EP-A- 0220273 JP-T- 62502617 AU-B- 571128 | 20-11-86 06-11-86 16-10-86 18-11-86 06-05-87 08-10-87 31-03-88 |
| EP-A- 0199301 | 29-10-86 | AU-A- 5636386 JP-A- 62012799 | 23-10-86 21-01-87 |
| EP-A- 0187041 | 09-07-86 | JP-A- 61233700 | 17-10-86 |
| WO-A- 8702038 | 09-04-87 | AU-A- 6299286 BE-A- 905492 GB-A- 2181435 FR-A- 2587720 SE-A- 8604007 NL-A- 8602422 FR-A- 2593519 JP-A- 63068075 | 09-04-87 25-03-87 23-04-87 27-03-87 26-03-87 16-04-87 31-07-87 26-03-88 |
| FR-A- 2593189 | 24-07-87 | WO-A- 8704459 AU-A- 6891187 EP-A- 0239425 | 30-07-87 14-08-87 30-09-87 |

EPO FORM P0472

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No.12/82